

How to live with cheap oil

The much-reduced price of oil must be an economic benefit in oil-consuming economies, but it is not too soon to begin preparing for the next price rise.

THE troubles caused by the declining price of oil continue, which must be the only satisfaction the impoverished oil producers derive from their inability to make the cartel called OPEC (for Organization of Petroleum Exporting Countries) effective. In Britain, the spot market for crude oil is in chaos, as traders have sought to back away from contracts signed when prices of \$15 per barrel seemed a very long way off. In France as well as Britain, the nationalized coal-producing industries are having to reduce their prices to remain competitive, with the result that they will be asking the governments that keep them going for larger subsidies. In India, the prime minister is in trouble for having put up the price of gasoline when everybody expects prices to be falling.

In both Britain and the United States, the falling price of oil has implications for taxation policy. The British government is saying that reduced revenues from North Sea oil will prevent it from reducing personal taxes in next month's budget (but that may be a blind). In the United States, falling prices for oil imports (now roughly a third of total consumption) have brought back into fashion the notion of an import levy (called a "fee" so that nobody should think the administration is increasing taxes). The domestic producers would welcome the protection, while the government cannot be indifferent to a scheme that might trim the federal deficit by as much as 10 per cent. But some countries dependent on oil revenues, Mexico in particular, will need financial help of a kind not now in prospect.

Where will all this turmoil end? The surprise is the speed with which the price of oil has fallen. At the time of OPEC's last formal meeting in December, when Saudi Arabia made it plain that it would not thereafter stint on oil production, the price was certain to fall below \$25 per barrel, but \$15 appeared beyond the bounds of possibility (see *Nature* 319, 2; 1986). Yet the slump should not have been the bombshell that it is. Over periods of time measured in months, the demand for oil and its products is inflexible. If the supply should suddenly increase, as it may have done to the tune of more than the equivalent of the 100 million tonnes produced each year in the North Sea, the oil producers are bound to be in a fix. Those with oil costing more to lift out of the ground than the going market rate will of course shut down their production wells, but others will have an incentive to keep on producing to generate the cash to amortize the capital cost of their facilities, almost always a larger component than the mere cost of recovery in the ultimate price. So, in the short run, the lower bound to the price of oil will be that at which enough producers with high recovery costs go out of business to balance the increased supply. Even in high-cost regions such as Alaska and the North Sea, there is probably some way to go before that point is reached.

Oscillation

Long before then, other forces will tend to stiffen the price of oil. Low-cost producers become high-cost producers when their wells begin to run dry, making secondary recovery necessary. And although it is unlikely that people will drive their motor cars further or more furiously simply because the price of gasoline has fallen, the industrial users of energy who have been switching

from oil to coal in the past few years will begin to switch back again. On a similar timescale, the oil producers will go slowly in the development of known resources, and will cut back on the exploration for new fields. Much of that is happening already.

Meanwhile, at any point, Saudi Arabia will be able to change the rules by which other oil producers must live simply by adjusting the output from its unique resource, the world's largest known reserve of cheap petroleum, destined to outlast any other comparable reserve at any imaginable production rate. There can be no complaint if Saudi Arabia chooses to follow such a strategy, but the consequences would be a cyclical fluctuation of the price, which may be less comfortable for the oil consumers than a high but stable price.

How should the oil-consumers prepare for such a strategy? One of the most striking features of the past decade of high oil prices is the way in which industrial economies have been compelled to be more economical in the use of energy. In the United States, for example, where Gross Domestic Product has increased by nearly a third since the first sharp price increases in 1973, oil consumption has actually fallen by roughly 10 per cent. In Western Europe, always more economical in the use of energy (by the yardstick of production), much the same has happened. If low prices had come to stay, it would make sense that part of the benefit should accrue to oil consumers in the generally reduced cost of relatively inefficient equipment. But since the price of oil may fluctuate on timescales measured in months, while decisions made now about energy-using plant will exert their influence in energy efficiency on timescales measured in decades, there is a strategic case that industrial economies should bias their systems to preserve some of the economic adaptations forced on them in the past decade.

But even the efficient use of energy entails an economic cost, that represented by the usually greater cost of relatively efficient equipment. If industrial economies were to follow the strategy of pretending that the cost of oil is still, say, \$25 a barrel, which might be done by sufficiently great import levies, people buying equipment would make the same decisions this year as last, but the governments concerned would have much increased revenues that might then be distributed by means of reduced taxes to individual oil consumers. The new benefit to the economies concerned would then be measured by the difference between the tax forgone and the continued extra cost of energy-efficient equipment. There would be some immediate economic benefit, but less than if the fall in the price of oil were fully reflected in domestic prices. As it happens, the proposed levy on oil imports into the United States would be a step along that road, but one whose size (\$8 a barrel is talked about) appears to be entirely arbitrary, or at best determined by the price that will keep domestic producers happy.

Other industrial economies are less well placed than the United States, locked as they are into competition with each other and even more eager that the falling price of oil should be made to signal the end of the long recession. Even without an international agreement on oil price strategy, there is much to be said for policies like that proposed for the United States (provided the levy is redistributed as reduced tax). □



Planning globally

There are plans for an international global research programme; its attainment will be difficult.

EIGHTEEN months ago, the general assembly of the International Council of Scientific Unions (ICSU) at Ottawa considered several new schemes for collaborative research, of which the best-grounded was that for a study of phenomena that are intrinsically global. Among the cheer-leaders at Ottawa were Dr Herbert Friedman (best-known for having found celestial X-ray sources by the use of captured V-2 rockets) and Professor J.A. Eddy, who rediscovered the Maunder sunspot minimum of the seventeenth century. They have now persuaded the US National Academy of Sciences to publish* what amounts to a prospectus for their study, which is both interesting and literate.

The starting point is the International Geophysical Year, now remembered fondly not merely as a proof that collaboration can succeed but also as a time when governments would spend money on it. Those advocating the new study have ingeniously called it the International Geosphere-Biosphere Programme, which has the initials IGBP. The project is a programme in search of sponsors.

The case for doing something is compelling. The artfulness of the document, which has been prepared by the US committee for an IGBP (note the indefinite article) of which Eddy is the chairman, is that it concentrates on the unsolved problems that abound. Why is there so much less water in the stratosphere than one would expect? What happens to the carbon dioxide released from the surface of the Earth, from wood and forest fires as well as fossil fuel, that seems not to accumulate in the atmosphere even when allowance is made for solution in the oceans and for increasing biomass? The problem is partly the familiar difficulty of subtracting from each other two large uncertain numbers (for net release and net absorption by known sinks), but there must be some undiscovered sink to explain why as much as 2,000 million tonnes of carbon dioxide is lost each year. Then, separately, where does all the methane in the atmosphere come from? What, after decades of speculation on the subject, is the relationship (if any) between solar activity and climate?

That a coordinated attack on such questions would be worthwhile is hardly in dispute. The sales pitch judiciously steers away from some of the more obvious traps, that the programme might become stuck with acid rain to the exclusion of all else, for example. It is to be hoped that the document will be read sympathetically by those who will have to pay for it.

Why should they bother? It is not necessary to be fixated by the prospect of environmental doomsday to acknowledge that even the natural fluctuations of the physical environment revealed by the historical record are potentially, to make the very least of them, a great inconvenience to governments. The prospect of another glaciation in some thousands of years may be too distant to evoke budgets on the scale required, but mean sea-level changes such as those now apparently under way could easily cause trouble on a timescale not very different from that on which countermeasures would have to be planned. But the essence of the case for IGBP is that without an understanding of the natural fluctuations of the physical environment, we are all vulnerable to the unexpected.

So how should an IGBP be organized? The prospectus says that the new programme would coexist with international programmes already in being, the World Climate Research Programme for example, and would in no sense subsume them, which makes sense. But what is proposed would itself be a research programme with a timescale not very different from some of the global problems it would tackle. It would depend for its success on a continued and even accelerated flow of data from sources already active in the field, many of them controlled by governments. So an IGBP would have to be a long-term project, and one that was able effectively to stimulate research as well as

helping to synthesize the data produced by others. A permanent international institute would almost certainly be less competent than existing laboratories working in parts of this huge field.

So why not aim at a radically different kind of international organization, one equipped with a small core of people with time to think and with funds sufficient to function internationally, as a foundation, making grants to people able and willing to fill obvious gaps in present understanding? The World Health Organisation's Special Programme of Research and Training in Tropical Diseases should be an encouraging model. The plain truth is that a project that is global in its essence deserves to be supported in a manner that is also global. □

**Global Range in the Geosphere-Biosphere* (National Academy Press, 1986).

Overheads undermined

The US government should think again about its scheme for cutting the cost of research grants.

THE research universities in the United States resist the use of the term "overhead" to describe the sums they receive from grant-making agencies in addition to the direct costs of successful research grant applications. The reason for this diffidence is obscure, but appears to derive from institutions' insistence that the "indirect costs" are real costs, those involved not merely in the administration of research grants but in the provision of that minimal support for research groups that will allow them to succeed in an increasingly fierce competition. But there are also suspicions that some universities fear that calling indirect costs overheads would lead to the fixing of uniform rate for all universities receiving grants instead of the present system in which different universities are differently rewarded.

This background has a bearing on the threat that now confronts the hundred or so research universities, the proposal of the Office of Management and Budget to amend the much derided circular A21 which regulates the way in which direct costs are accounted for and indirect costs determined. The intention is to fix a flat rate for the payment of indirect costs equal to the mean of the rates fixed in the past for different universities. Under the new arrangement, some will be better off, but because of a familiar statistical circumstance, most will be worse off. The explanation is merely that the universities traditionally more successful in the competition for research funds have also been those with the highest rate of reimbursement of indirect costs. Indeed, according to the draft of the amended regulation now published in the *Federal Register*, the federal government hopes to save about \$100 million a year by this device.

University presidents are already up in arms about the issue, and rightly. Their circumstances are very different from those elsewhere, in most West European countries and even Japan, where governments recognize an obligation to support the infrastructure of research at universities. In the United States, state governments have in the past recognized such an obligation, but there is no way in which the important private universities can be helped, unless from a levy on student's fees, except by federal support.

What is to be done? Universities will not serve their own interests if they protest too vigorously at the principle of uniformity, which offers at least an escape from the separate haggling over what the rate should be at different institutions. It would even be sensible that the universities now hoping to get the amendment changed in draft should offer to transfer from the direct to the indirect category of allowable costs some of the incidental spending (on travelling and publication, for example), which at present clutters up most research grant applications. But then the universities should insist that the rate which has been set is far too low. In Britain, the government reckons that 30 per cent of the recurrent cost of universities is for research (which amounts to roughly £600 million), while the direct cost of research grants to universities is less than half as much. □

US shuttle

More changes in NASA's higher echelons

Washington

BEFORE the resumption of the public inquiry by the President's commission at the beginning of this week, the US National Aeronautics and Space Administration (NASA) had begun clearing the ground for a reorganization at the top. The post of head of the shuttle programme has been filled by the appointment of Rear-Admiral Richard H. Truly, a former astronaut. His predecessor, Jesse Moore, who had been appointed at the end of last year as director of the Johnson Space Center at Houston, has taken up that post sooner than planned.

Meanwhile, the administrator of NASA, James C. Beggs, remains on indefinite voluntary suspension, after his indictment at the end of last year on criminal charges levelled by the Department of Defense against General Dynamics, the military contractor of which Mr Beggs had previously been a vice-president. NASA is for the time being the responsibility of William Graham, previously Beggs' deputy, but the appointment of a permanent administrator is expected shortly.

The commission's preliminary inquiries have taken its members separately to sites such as Cape Canaveral and the Morton Thiokol plant in Utah at which the suspect solid-fuel boosters for the shuttle were manufactured.

In the process, Dr Richard Feynman, the Nobel physicist from the California Institute of Technology, seems on the way to becoming as well known to the general public as he is famous within the academic community. He spent much of last week at Cape Canaveral looking into reports that parts of the lost shuttle craft, and in particular the lower part of the booster that appears to have failed, became exceptionally cold in the hours before the launch. Several theories of the processes that might account for this have been developed.

Documents released by NASA support the speculation (see *Nature* 319, 609; 1986) that the fault responsible for the disaster on 28 January was the failure of one of the synthetic rubber-based sealing rings joining the lowest of the four sections of the right-hand booster rocket to that above.

At an earlier stage, there had been concern about the putty used for filling the joints. According to one of the documents, the original material (based on zinc chromate) had contained asbestos, but this was withdrawn from the market under Environmental Protection Agency regula-

tions. More recently, putty containing asbestos was obtained from a Canadian source, but is described as having been more "tacky" and thus less suitable. By August last year, Thiokol had submitted no fewer than 43 alternative designs for

the case-to-case sealing joints, one of which has been incorporated into a new design of rocket case. Estimates of the time needed to modify existing cases range from 3 to 12 months.

Fears that similar design problems may affect the solid-fuelled ICBM rockets of the US Strategic Air Command are apparently inapplicable. A spokesman at Omaha, Nebraska, said last week that both the Minuteman and the MX (Peacekeeper) rockets consist of three stages each of which is integral in construction. □

Solar telescope

US and China join with Europe

LEST, the Large European Solar Telescope project, scheduled to become operational at the beginning of the 1990s, has changed its name. At the fourth general assembly of the LEST foundation in Tenerife (29 January–3 February 1986), two new countries, China and the United States, formally joined LEST, and the "E" of the acronym was accordingly changed to mean "Earth-based". The participation of these two countries (and also of Australia) will undoubtedly affect the choice of a site for the telescope. The original plan was for a site in the Canary Islands, but, with the United States interest in the project, the idea of Hawaii has, in the words of one participant in the general assembly, "gradually crept in". Solar photographs from the new Swedish telescope at La Palma convinced many of the participants that an Earth-based facility makes scientific sense.

One main purpose of LEST will be to study solar magnetic fields (by means of polarized light observations), and, in particular, their evolution with time. This will require not only a good telescope (the designs for LEST show a 2.4-m aperture; the largest solar telescope at present is the 1.5-m instrument at Kitt Peak, Arizona), but also a site with good optical quality for a significant proportion of the time. Although La Palma can occasionally have very bad weather in winter this is of relatively short duration, and the photographs obtained by the Swedes since their telescope went into operation on 13 December greatly impressed the participants in the LEST assembly. Dr Hou Ning-sheng, of the Yunnan Observatory in Chi-

na, was particularly enthusiastic about the Swedish results, which clearly show that time-development studies are feasible from a ground-based instrument.

Sweden has a special interest in LEST. The Royal Swedish Academy of Sciences has acted as "host" to the LEST organization since it was established in 1983, and has provided seed-money for feasibility studies. Dr Kai-Inge Hillerud of the Swedish academy is the executive secretary of the organization. The project manager and newly appointed director of the project is a Norwegian, Dr Oddbjørn Engvold of the Theoretical Astrophysics Institute in Oslo.

Apart from the choice of site (for which testing is about to begin in the Canaries and Hawaii), the two main problems are finance and technology. There are now nine nations in LEST (Sweden, Norway, West Germany, Switzerland, Italy, Israel, Australia, China and the United States), each represented through a non-governmental organization. How the cost of the project (estimated at US\$20 million) will be shared has not yet been decided. One proposal is for minimum contributions based on the ratios of members' Gross National Products, possibly with additional support from foundations and individual benefactors. Some private finance has, indeed, already been forthcoming — the technical design study for the proposed telescope was financed by Robert Maxwell, the British publishing magnate.

The "next generation" technology envisaged by this study will include active and adaptive optical systems to correct for atmospheric distortion and a sophisticated data storage and computer base. An important feature of LEST will be that observations can be carried out either by on-site staff or from remote-control observation rooms in the host countries. Although Earth-based, LEST will, in effect, be run, in Dr Engvold's words, as a "space instrument" and the technological problems it will pose are likely to prove extremely challenging. **Vera Rich**

CERN inquiry

AN international review of the European Organization for Nuclear Research (CERN) is to be set up following a unanimous decision by its council last week. The result of a British initiative (see *Nature* 319, 525; 1986), the review is expected to be completed within twelve months. □

Spanish universities

New law leaves unsolved problems

Barcelona

ALTHOUGH Spanish universities will benefit in many ways from the new law now coming into effect, many academics believe the legislation is too little, too late, or both. Even so, the university reform law, passed in 1983, is being implemented methodically; a new university structure should be in place by 1988.

During this month, most universities have been electing their academic heads (presidents in US universities, vice-chancellors in Britain), often with surprising results. Those standing for office have mostly been professors with a strong political background. The chief candidates for the post at the University of Barcelona were two members of the Catalan Parliament; after a close fight, Professor Bricall, an independent economist and a member of the socialist parliamentary group, defeated a communist professor of Greek. At the Polytechnic University of Catalonia, on the other hand, a member of CDC (the nationalist Catalan party) defeated a candidate supported by socialist and communist professors.

Other Spanish universities have also found the elections to have had a political dimension, in part because of the political attention paid to university reform during the past few years, while the long overdue reform law was being hatched. Among the important changes are the decision that future appointments will be made more logically than under the previous system of *oposiciones*. In the words of Sr C. Virgili, "candidates will be judged by what they know and not what they are ignorant of".

At the same time, the autonomy of the universities is being strengthened after a long period (since the civil war) when they have been in a straitjacket, even to the extent that universities will be able to determine their own curriculum. Similarly, courses of study leading to the PhD degree are being consolidated, which will enhance their status. The present socialist government is providing strong financial support for the reforms and has enormously increased its support for university research.

Indeed, the Spanish parliament is about to approve a new package of legislation dealing with science, which is eagerly awaited and which will transform the administration of science in Spain. The law has been delayed by changes demanded by the Senate, but should be complete next month.

The reasons why not everybody is satisfied with the reforms go back for many years. Especially towards the end of the Franco regime, political activities tended to dominate the affairs of the universities;

afterwards, despite active discussion of the need for university reform, the early democratic governments were unable to carry through legislation. And even now, some universities appear to provide a playground for future politicians or even a shelter for those who have failed in the parliamentary elections and are waiting for better times.

Moderate discontent is illustrated by the views of Professor A. Nieto, a former president of the Spanish research council, who says in article published by the March Foundation (*Cuadernos Universitarios*, no. 225) that the new law will not by itself solve the problems facing the Spanish universities. When the universities are governed by the rules of the parliamentary government, Nieto says, people are compelled to fight for power and to disregard the essential functions of the universities. What matters in academic life is therefore the number of votes a person controls, not his performance in teaching and research.

Nieto argues that a university may be governed either by the state or by society, but "never by its own professors, who should be the servants of the university, nor by its students, who are its beneficiaries". On autonomy, he says that no-

body knows what the effects would be because Spanish universities have never been, and are not now, autonomous, but he agrees that more autonomy might lead to "internal qualitative differentiation" within universities, with the result that some would excel over others.

As things are, the role of university professors as academic leaders is ambiguous and confused. For example, as the reform law is applied, the constitution of the committees appointed to fill vacant faculty positions enormously favours inbreeding; the students of the committee chairman are the most likely to succeed. Many academics fear that this state of affairs will hamper the improvement of research activity. Another worrying sign is the large number of the sons and daughters of middle-class families who move to the United States for their undergraduate studies. Has the turbulence of the 1970s undermined people's faith in the universities?

Despondently, Nieto believes that change will come only from outside, "as a consequence of a profound change in the social and political mentality" of Spain, which he believes will not come quickly. But others recognize that the new law is a legal framework for reform, and look to Spain's accession to the European Community from the beginning of this year with great hope.

Juan A. Subirana

Hydroengineering

More controversy over Danube

THE controversy over the Gabčíkovo-Nagymaros hydroelectric project on the Danube (see *Nature* 317, 9; 1985) continues to acquire international overtones. Since the Austrian government last year undertook to finance the construction work at Nagymaros in Hungary, the Austrian "greens" excused their government of "exporting" its environmental problems to Hungary. The Austrian greens therefore decided, according to their leader, Dr Gunther Nenning, to "export democratic, non-violent constitutional resistance" to Hungary. The first such attempt, a "protest walk" through Budapest on 8 February, was broken up by the police, but the greens, backed by their West German counterparts, seem determined to keep up the pressure, and hope to make Austrian involvement in Nagymaros a major issue in the forthcoming Austrian general election.

At the same time, attempts by the unofficial "Danube Circle" of Hungarian environmentalists to finance an "independent" survey of the possible consequences of the dam are attracting considerable attention in Sweden. The money for the survey would come from the "Right Livelihood Award", presented annually

by the Swedish parliament. The awards are conferred not on individuals but on groups. Since the Danube Circle is not recognized by the Hungarian government, it cannot import the prize money. Dr Janos Vargha, who represented the circle at the prize ceremony last December, has indicated that the members would pay any organization competent to carry out the survey, provided the survey would be independent. Representatives of the Right Livelihood Foundation have begun lobbying international figures for the Danube Circle to be allowed to use the money as they wish.

Western interest in the dam has not gone unmarked in the Comecon bloc. The project has been praised both as a joint Czechoslovak-Hungarian venture and as a major achievement of socialist cooperation. Officially, all environmental hazards of the dam have been foreseen and well-provided for by surveys which include not only the Hungarian and Czechoslovak academics, but Soviet experts as well. This is the first time that direct Soviet involvement in the plans for the dam has been openly mentioned, and it has been widely interpreted as a direct response to growing Western interest.

Vera Rich

Environmental pollution

Does Japan have acid rain?

Tokyo

SUDDEN fears that Japan, like Europe and North America, may be suffering from the effects of acid rain have triggered the start of an emergency survey of the forests around Tokyo. The Environment Agency announced the survey last week and preliminary data should be collected by April.

The emergency survey was provoked by two new research reports, from the Gunma Prefectural Health and Pollution Research Centre and Chiba University, that show evidence of extensive damage to trees by air pollution in the Kanto plain (the plain surrounding Tokyo). The news came as a shock; in Japan rainfall is abundant and rivers short so pollutants should be quickly washed away. But what was thought more important is that Japan has some of the world's strictest air pollution legislation.

The legislation was occasioned by some appalling cases of air pollution during Japan's period of rapid industrial growth. In the 1960s, as the Japanese economy



boomed, oil-fired power stations were introduced very rapidly and sulphur oxides were poured into the air. In the refinery town of Yokkaichi, for example, conditions grew so bad that special face masks had to be issued to primary school children. Public protests led to government action: by the end of the 1960s, measures were under way to desulphurize heavy oil and to reduce the sulphur in emissions from smoke stacks. Then in the 1970s, a fierce battle was fought over another major air pollution source: automobile exhaust emissions. Despite delays, the citizens groups and local governments that took on the giant automobile manufacturers won. Emission standards for carbon monoxide and nitrates were adopted which were the strictest in the world. Petrol has been lead-free for more than 10 years.

Measurements of air pollution seemed to confirm that speedy action had prevented the occurrence of acid rain. Unofficial surveys carried out by volunteer groups and coordinated by scientists from the Environment Agency's research institute suggest that the annual mean pH of

rain in Japan is in the range 4.3–4.4 with a minimum annual average of pH 4.0. As unpolluted rain has a pH around 5 from natural inputs of sulphur oxides (from volcanic eruptions and the like), these levels of acidity seemed no cause for concern. Even during the rainy season from June to July when frequent drizzle favours the production of acid rain, pH levels did not appear particularly low. A survey of the area around Tokyo Bay carried out in the rainy seasons of 1975–1978 showed an average pH of 4.4 for Tokyo, and of 4.0 and 3.8 for the heavy industry areas of Kawasaki and Yokohama, respectively.

Despite these generally optimistic findings, there have been occasions when rainfall has become very acid — around pH 2.8 — and caused skin and eye irritation, but these were rare events related to unusual weather conditions. Observations of air pollution also remain patchy and insufficient funds are available to complete the analysis and publication of the most up-to-date data collected.

And of course the biggest problem of all is that without clear understanding of the mechanism of formation and effects of acid rain it is very hard to know exactly what should be monitored. Ozone levels, for example, are known to be high in Japan and may be an important factor in forest damage. And although the pH levels measured around Tokyo are higher than those found experimentally to cause damage to plants, they are not much different from those in areas of Europe affected by "acid rain".

New results described at a small conference held at the National Institute for Environmental Studies at the end of January came as something of a shock. A survey of *sugi* (the Japanese cedar) carried out by Dr Keiji Takahashi of Chiba University in the Kanto plain revealed much damage that could not be attributed to normal environmental causes, salt-winds, insects, winter winds and the like. *Sugi* is the most widely planted forest tree in Japan and because it is often grown in shrine precincts there are many old trees available for study. Many trees showed symptoms of "die-back" — loss of leaves on the top-most branches apparently similar to that seen in the dying forests of northern Europe, with old trees most affected. At the end of last year, worrying results were also announced by Dr Kyoichi Sekiguchi of the Health and Pollution Research Institute of Gunma Prefecture, to the north-west, and downwind, from Tokyo. In some parts of the north-western Kanto plain, as many as 80 per cent of cedars seemed to be affected. In Maebashi City, some 100 km north-east of the main industrial zone surrounding Tokyo Bay, sul-

phates and nitrates are transported by wind and in the rainy season produce rain with a pH around 3.5.

The new survey will attempt to find out whether a real problem exists. A budget of 8.2 million yen (US\$45,000) has been allocated for study of an area within a hundred kilometres of Tokyo which includes both the Kanto plain and the edges of the mountains that surround it. The Forest Agency's research institute will carry out survey of the cedars while the Environment Agency will look at soil, water and air pollutants. If there is cause for concern then it is likely that larger projects will quickly be launched. **Alun Anderson**

The Japan Prize

Tokyo

WINNERS of Japan's answer to the Nobel Prize were announced in Tokyo last week. This year Japan Prizes go to Professors Willem J. Kolff (University of Utah) and David Turnbull (Harvard University).

The prize, now in its second year, is worth 50 million yen (more than US\$250,000) and is to be awarded to two people each year. It was set up with funds provided by Kohnosuke Matsushita, head of the giant electronics company that bears his name, and is run by a foundation under the guidance of the Prime Minister's office. Its aims are a little different from the Nobel prize, for it is intended to reward achievements in science and technology that contribute to the peace and prosperity of mankind: thus, appropriately enough for Japan, it has a slightly more applied-science flavour than the Nobel prize.

Two fields are selected each year. This year, as last, both are fields in which Japanese industrial researchers are active and in which the Ministry of International Trade and Industry is supporting large projects, although, perhaps out of modesty, a Japanese scientist has yet to win the prize. Professor Willem Kolff was chosen for his contributions to "medical technology". Professor Kolff is known as the "father" of artificial organs. He developed an artificial kidney in the Netherlands in 1943; after his move to the United States, he succeeded with an artificial lung and has played a key role in research on artificial hearts. He is now a US citizen.

Professor David Turnbull gained the prize in "material science" for his work on the theory of amorphous materials. In the 1950s he predicted that any liquid, if cooled quickly and if sufficiently viscous, can be transformed to a "glassy" amorphous state. His hypothesis has since been experimentally confirmed and has led to understanding of the structure and properties of numerous glassy metals, alloys, polymers and ceramics. These new materials are finding diverse applications, from solar cells to lightweight but strong structures. **Alun Anderson**

Museums

Expenditure cut to the bone

AS EVERY London schoolchild knows, dinosaurs, whales and even computer games are to be sampled for free at the Natural History Museum in South Kensington. But if BM(NH) — the British Museum (Natural History), which comprises the Natural History Museum, the Geological Museum also in South Kensington and the Zoological Museum in Tring — cannot gain considerably more funds from government or by patronage, free admission to its museums is likely to be abolished, scientific services charged for and scientific staff reduced by at least 10 per cent. Such is the conclusion of BM(NH)'s first corporate plan, announced this week.

In formulating its corporate plan, BM(NH) has estimated that the gap between its predicted expenditure and income is likely to rise from £1.2 million next year to £2.7 million by the end of the decade. Cutting running expenses and operating a policy of not replacing all retiring staff will meet a quarter of this. By spending more on marketing, a net gain is forecast in income from publications, catering and sponsorship. And a little further income will come from introducing charges for scientific advice, loans and the use of the museum's collections. Even so, there remains a considerable gap which will need to be met by staff cuts, admission charges or both.

In one budget-balancing forecast, in which admission charges are introduced in 1987–88 and yield £1.6 million by 1990–91, the staff of the museum are reduced by 35 to 805 in the same period, with the brunt of the cuts falling on the scientific staff. They numbered 361 in 1983–84, will be 335 this year and are forecast to be 300 in 1990–91. "The redundancy of expert staff ... will be implemented only if totally unavoidable", states the plan.

Asked why it should be the scientific staff that suffer disproportionately, R. Saunders, secretary of BM(NH), said that is only the exhibition side of the museum that can expect to attract sponsorship so its staff must be maintained. Only the government and its agencies can maintain the level of scientific staff.

If admission charges are introduced, it is not clear whether they will be compulsory or voluntary, as at the Metropolitan Museum in New York and the Victoria and Albert Museum in London. The latter is already finding a considerable decrease in visitors despite the non-compulsory charges; how London's schoolchildren and the Natural History Museum will come to terms with charges and their effects is unclear.

Peter Newmark

US universities

New medical school planned

Washington

THE Howard Hughes Medical Institute (HHMI) and Yale School of Medicine last week announced plans to build a joint centre for molecular medicine worth at least \$25 million on the medical school's campus. The facility may provide a model for future medical schools by crossing departmental boundaries, which many researchers find increasingly cumbersome and artificial. Its select research contingent may also have to grapple with the jealousy and elitism that have plagued some of Hughes's 22 academic collaborations in the past. But such shadows are not dimming the expectations of the centre's planners, or of the few investigators who have already won places in the facility in 1989, when its doors will be opened.

The partnership venture is an extension of a relationship with Hughes that Yale has enjoyed since 1977, and follows close on the heels of the \$5,000 million sale of Hughes Aircraft Company that made HHMI the wealthiest philanthropy in the United States (*Nature* 315, 531; 1985).

Construction of the 110,000-sq. ft building will commence within 17 months with \$15 million from Hughes and with Yale making up the difference. Each of the centre's four floors will be devoted to a particular programme of study and three have already been chosen. Hughes will underwrite the molecular neurosciences and molecular genetics divisions, while Yale will manage a molecular oncology programme and the fourth programme, for which neurobiology, immunology and receptor biology are among the main contenders.

The centre will also house four or five "core facility" laboratories that will provide services such as protein sequencing, peptide and oligonucleotide synthesis and monoclonal antibody production for the centre staff and all the medical school's laboratories.

Joint Yale/HHMI search committees will recruit between 25 and 30 investigators to fill the facility. Yale hopes the centre will draw expertise in areas such as oncology in which its own faculty is a little lean. But some Yale researchers are concerned that collaboration might rank among the "monopolies of excellence" which have been accused of draining talent and recognition from less conspicuous, but equally deserving, universities. And while the scenario of Hughes investigators and Yale faculty working side by side seems at first idyllic, some anticipate personal friction of the kind that frequently attends cooperation between more and less generously endowed teams. The geneticists and neuroscientists working on Hughes floors will be afforded the luxury

of long-term goals and spending flexibility, while Yale researchers will continue to toil in traditional competition for funds.

Finally, the notion of interdisciplinary programmes, however appealing, may create havoc in the planning stages. What Dean Leon Rosenberg calls "a melting pot of faculty" could turn into a pressure cooker if departmental loyalties prove not so superfluous.

None of these reservations, however, are likely to prompt rejection letters from those invited to join the centre's group. For biomedical research, says one Yale professor, the bottom line on the centre for molecular medicine is "wonderful".

Karen Wright

Biotechnology

Will Pharmacia

Washington

SURPRISING revelations about Refaat El-Sayed, president of Swedish biotechnology company Fermenta AB, have slashed the company's share prices and put on ice its plans to buy into two pharmaceutical concerns. The man who increased Fermenta's earnings by SKr 1,500 million in three years resigned last week following his confession on 13 February to having lied about his academic credentials. Then El-Sayed capped the controversy with a disclosure that he had bought 4 million Fermenta shares shortly before releasing 1985 profit and loss figures, a move that may violate laws set down by the Stockholm stock exchange.

El-Sayed claims he bought the shares in connection with a tentative agreement with Volvo announced early in January, by which Fermenta would acquire the Swedish conglomerate Sonesson and Volvo's 40 per cent stake in Pharmacia, the second largest pharmaceuticals company in Sweden.

A tender offer was set stipulating a Fermenta share price of SKr 220. Shares soared over SKr 300 after the announcement, but plunged to about SKr 150 when El-Sayed's credibility was called into question. Now Volvo and Fermenta are renegotiating the deal; meanwhile, Volvo is denying any involvement in El-Sayed's purchase.

Pharmacia, the coveted object of the transaction, is preserving a "business as usual" attitude until a bargain is finally struck — or dissolved — between the two companies. Although Pharmacia will be listed independently on the Stockholm stock exchange regardless of the outcome of the negotiations, analysts say its executives were facing the acquisition with some

Plant breeding

The seeds of commerce

ANOTHER firm step has been taken by the UK government towards privatizing the National Seed and Development Organisation (NSDO), which produces and markets seeds of new plants that emerge from government-supported research, and the major source of NSDO's new plants, the Plant Breeding Institute (PBI) of the Agricultural and Food Research Council (AFRC). Within hours of last week's announcement that the government is to seek financial advice on how best to proceed with the sale of NSDO and the plant breeding, as opposed to the plant science, programme of research at PBI, one company was singing its own praises as the best possible buyer.

NSDO has been a candidate for priva-

tization for nearly two years because it is a profitable business. Last year it recorded a profit of about £7 million, of which £4.65 million was returned to the Treasury, while only about £2.5 million was spent by the council on plant breeding research at PBI, from which nearly all of NSDO's new varieties have come.

But the plan to make the sale more attractive by including half of PBI is more recent and has been opposed by the institute, which fears that the remaining plant science research will suffer as a result. Reacting to the decision the institute's director, Professor Peter Day, said that the institute has always prided itself on covering the spectrum from molecular biology to the production of finished new varieties of plants and the challenge now will be to maintain the status of the science programme on its own. In any case, the institute will now have to develop a new strategy of research, probably concentrating more on alternative crops than on the mainstays of NSDO and almost certainly acquiring some breeding facilities.

Because financial cuts have made it impossible to maintain the *status quo* at PBI, and because the case for supporting research there by the profits of NSDO has failed to win the day with a government set on privatization, Day will now have to work out a new strategy for the remaining half of the institute. The worst outcome, he said, would be if the science and breeding activities were altogether separated as might happen if the remaining scientists are moved elsewhere. The best that can be hoped for, says Day, is that business will continue much as now except for the inevitable differences necessitated by the terms of any formal relationships between the remaining scientists and the new employers of its plant breeders, or at least those that are lucky enough to gain such employment.

Such might be the case if the Agricultural Genetics Company of Cambridge became the new owner, according to Roger Gilmour, its chief executive. The company was set up three years ago with the help of the government-owned British Technology Group and was given first rights to develop and market discoveries made by plant scientists at AFRC institutes. The rights are for five years in the first place and negotiations for their renewal are due this summer. At least as important to the company as the renewal of these rights is that it should continue to have access to a plant breeding facility and a seed marketing organization. The company's future seems bleak if it fails in its bid to buy NSDO and the part of PBI that will be for sale.

When the sale will be and exactly what

will be on offer is anyone's guess. It seems very likely, however, that the two places on offer will be sold together and to a single buyer.

The government is obliged to go through a procedure of appointing an independent financial adviser to report on the best way to go about the sale. Only when such an adviser has been chosen from applicants and has reported to the government is any further statement expected. It may then also become clear whether it is the Treasury or AFRC that will be the beneficiary of the sale.

Peter Newmark

Ariane satellites up

FRENCH — and Swedish — space scientists breathed sighs of relief last week as the mostly French-built European space launcher Ariane lifted two satellites safely into orbit. This was the first launch for some months, after the last was aborted when a third stage rocket motor failed. The present launch, which put up the first French Earth-observation satellite, SPOT1, and the Swedish research satellite Viking, had been delayed when a German-built cooling system was found to be leaking.

SPOT1 is following a near-polar orbit at an altitude of 822 km. It carries both vertical and sideways-looking cameras, so the same site can be pictured twice from different directions on successive orbits, as the Earth turns beneath the satellite's path, thus yielding three-dimensional images of the surface which will be useful for the mapping of remote areas. Some potential users in Africa, however, have complained that the technologies employed by SPOT1 will not be of much use for crop management and agricultural planning, but are much better tuned to monitoring defence installations and troop movements. As SPOT1 images will be available commercially, this could lead to some political difficulties.

Robert Walgate

Vera Rich adds: The Swedish Viking satellite will study the region of acceleration of incoming solar wind particles, which, when entering on the dark side of the Earth, become pumped in energy and in latitudes 65–75° interact with the atmosphere to produce the visible aurora. The Viking satellite has 40-m wire booms to trap the particles, and first results are expected in about two weeks.

The Swedes, however, are not totally committed to Ariane as a launch vehicle. The Swedish Space Corporation is considering production of a small (100kg) telecommunications satellite MAILSTAR, which will carry telex and telephone traffic to the third world. A tentative agreement has just been signed giving China first option as provider of launch facilities, should MAILSTAR materialize. □

Join Fermenta?

trepidation. Whereas Volvo has maintained a reputation as a "hands off" shareholder, El-Sayed is considered more likely to meddle. If the initial arrangement with Volvo were to be finalized, the Egyptian-born Swede would wind up with almost 50 per cent of Pharmacia's voting shares.

It is hard to say how closely El-Sayed's credibility is linked with shareholder confidence, but one analyst says that the man named 1985 Swede of the Year is now "lethal to his own company". El-Sayed owns more than 70 per cent of Fermenta.

The preliminary results presented last week, intended to restore investor optimism, contained profit figures of SKr 3,320 million that were augmented by minority interest from other Fermenta holdings, primarily the Italian pharmaceuticals company Peirell.

In the wake of these blemishes, says Fermenta, El-Sayed gave up his leading position in order to devote more time to the company's future, as director of long-range planning and acquisition. Others suggest that he acted on the insistence of Svenska Handelsbanken, Fermenta's parent bank, which had unwittingly proclaimed El-Sayed's credentials in a prospectus issued last summer for a SKr 210 million private offering in London.

Even if El-Sayed and Fermenta escape unscathed from the past week's turbulence, they are still in hot water with the Swedish environmentalist who incited it. Bjorn Gillberg at the environmental centre in Uppsala, who knew El-Sayed as an undergraduate at the town's university, first contested his claims to a master's and PhD degrees. But Gillberg's real concern is Lake Malaren which he believes is being polluted by a Fermenta fermentation plant.

Karen Wright

US science spending

United front on basic research

Washington

THE mutual congratulations of the National Science Foundation (NSF) and its authorization committee, the House of Representatives subcommittee on science, research and technology, at their first encounter of the budget season last week, were soured only by Representative Ted Wirth's (Republican, Colorado) attempt to find out why Colorado State University has not become a supercomputer centre. According to Erich Bloch, NSF's director, the explanation is simple: Colorado State was one of 22 competitors for the honour, but its application was simply not good enough to be one of the five chosen under Phase II of NSF's programme to build an academic supercomputer network. Bloch's scorn of the pork-barrel is well known, but the primness of his rejection of the taint last week may have been undiplomatically austere.

The issue of Colorado State's computer centre is a tiny cloud in what seems to be an otherwise clear sky. The university became part of Phase I of the programme in 1984, with the gift of a Cyber 205 machine from NSF, but this did not qualify it to become one of the five Phase II centres, on which NSF will be spending \$35.5 million next financial year if the budget survives. But is this not a waste of taxpayers' money? Wirth's simple question was simply answered with the information that supermachines are cheaper now than two years ago. Whether this will satisfy Colorado State and Wirth remains to be seen.

For the rest, only mutual uncertainty about this year's budget process, "meat cleaver" accountancy according to one committee member, disturbs the cosiness of the subcommittee's relationship with its only charge. But Bloch, careful as always, failed to seize the opportunity of saying which programmes NSF would have to butcher when the Gramm-Rudman axe falls compulsorily if the federal budget exceeds the legal limit later in the year. Bloch said only that NSF had been able to survive this year's cut of 4.3 per cent by selective economies in NSF's different programmes. This would not be possible if, later in the year, the Gramm-Rudman legislation should make cuts of "10, 15 or even 20 per cent" from next year's spending plans.

NSF is being dealt with even-handedly but not generously at this stage. The budget is due to rise by 8.4 per cent to a total of \$1,686 million in the financial year beginning on 1 October (see *Nature* 319, 526; 1986). Bloch pointed out last week that research and development supported by the federal government is due to increase even faster, with the result that basic research will be a shrinking pro-

portion of the cost of federally supported development (almost all of which is military). "The already small share of research dollars devoted to basic research is dropping."

The subcommittee and the agency it superintends appear this year to be more determined than ever to carry a torch for basic research. With economic competition internationally, persistent unemployment, budget deficits and a "faltering educational system", what could be more important than NSF and its support of basic research?" asked one committee member. Dr Roland W. Schmitt, the chairman of the National Science Board, NSF's policy-making partner, equally asserts that the academic system "is the basis of the economic defense" of the United States.

So how could NSF and the board remedy the "under-investment" in basic research in the United States? Both agreed that they cannot succeed alone. NSF spends less than a fifth of federal spending on basic research, which itself is only 14 per cent of government spending on research and development of all kinds. But Bloch and Schmitt agree on their objectives.

One need is to persuade other government agencies to pay more attention to

their own self-interest in the well-being of basic research. The condition of the universities, and especially of their equipment, needs attention. The declining numbers of US citizens going on to PhD courses is alarming. (Schmitt says that his experience, from a base at General Electric, is that many industrial laboratories would be "in deep trouble" if one half of the overseas PhD graduates did not stay on in the United States.) So more energetic steps must be taken to capture the enthusiasm of young people at an early stage, whence the renewed interest and spending of school curriculum development.

For the year immediately ahead, NSF plans to spend more on its academically based engineering centres, of which there are already six, with another four or five about to be announced and with funds for a further four in next year's budget. Extra funds will be found for biotechnology (in which there is now a plant sciences postdoctoral programme), computational science and "global geoscience" (where spending will be increased by \$18 million, compared with \$17 million in the present year).

That is what will happen if the budget survives. In the subcommittee's mood last week, it will not rob NSF of a single cent. But nobody can guess at this stage how far the committee's protection of its protégée will be undermined by Gramm-Rudman.

John Maddox

Japan likely to participate in SDI

Tokyo

THE Japanese government has been dropping broad hints that it will soon allow participation in the US Strategic Defense Initiative (SDI). Whether the hints are partly intended as trial balloons to see whether opposition remains is unclear, but they have been sufficient to draw a strongly worded warning from the official Soviet news agency Tass.

The government appears to have abandoned thoughts of participating directly in the SDI project through research in its own institutes. Instead, private industry will be allowed to accept contracts for SDI research. A final decision is likely to be announced before the annual summit of seven industrial nations in Tokyo in May.

A year has already passed since the United States invited Japan to participate in SDI, without the government having been able to come to a decision. Early on, obstacles were seen to participation in Japan's bans on weapons exports and on nuclear weapons research. But supporters of SDI within the government argue that joint research and development is different from arms exports and that even if nuclear energy is used in SDI, to pump X-ray lasers for example, that does not make them nuclear

weapons. Fear of arousing public opposition has, however, made the government very cautious. Two missions have visited the United States to seek further information and a third, which will contain business leaders as well as government representatives, is likely to be despatched in the near future.

Japanese corporations are mainly concerned that they might be left behind in the "technological dust" if they do not participate in SDI. Nor do they wish to miss access to the huge sums of money that are likely to be spent in the research phase of SDI. But there is also a strong awareness that Japan's own great economic success is not entirely unrelated to high expenditure on commercial development research and tiny expenditure on military research. For that reason, industry is keen to maintain a position in which the results of SDI research can be easily transferred to the commercial sector. But that position may not necessarily work to industry's advantage. Although under present Japanese law it would be hard to stop such transfer, it is likely that few contracts will come Japan's way unless legal protection of "military secrets" can be strengthened.

Alun Anderson

Causes of African famine

SIR—Brian Perry¹ stresses the need for constructive analysis and critical review of the causes of African famine, cites the opinion of Sinclair and Fryxell² that overgrazing around artificial water holes in semi-arid areas is a major cause of desertification and climatic deterioration and concludes that the improvement of the environment must involve not only technical, but also social and political factors. I believe that his analysis is correct, but fear that social and political considerations are likely to be more intractable than purely technical problems.

The root cause of the problem is ecological³ and can be overcome only by a far-reaching socio-economic revolution among the populations affected. The value systems of rural African communities are still largely those appropriate at the turn of the century, when most people were nomadic pastoralists or subsistence farmers practising shifting cultivation. Modern medicine and hygiene, and the suppression of tribal wars, led to great demographic change, with populations increasing fivefold or more in a few generations⁴.

The common feature of civilizations is not a value system but an administrative framework which allows the accumulation of agricultural surpluses and their distribution in time of famine, which can be viewed as behavioural adaptations to ecological pressures. A common factor has been the development of cities, which removed people from the land and formed a market for farmers, thus shifting the basis of the economy from subsistence to exchange.

Sufficient food for all the people of Africa could be produced by the practices already in use in those countries with developed agricultural sectors. The obstacle is the innate conservatism of the peasants, in contrast with fishermen in the area, who readily adopt new methods, probably because their product is perishable and must therefore be exchanged so that they are conceptually conditioned to a market economy⁵.

People will produce goods surplus to their own requirements only if they are suitably rewarded by either spiritual satisfaction, disposable income or the avoidance of punishment. The last is generally unacceptable under a liberal value system, which leaves income as the only practicable incentive. Unfortunately, money alone is not sufficient, for it has no value unless it can be converted into goods or services considered desirable by the recipient. The lack of opportunity for such trade is the major disincentive to rural development in Africa. It may be significant, in this context, that Malawi, which is surrounded by Mozambique, Tanzania

and Zambia, all of which have experienced shortages, has consistently produced a surplus of food. The difference is that in Malawi, education, which is universally perceived as desirable, is contingent on the payment of school fees. The provision of such services from government revenues alone is insufficient to ensure development of the economy.

The problem is thus socio-economic rather than technical, and it is in this area that research is most urgently needed. I feel that the key lies in the development of rural retail trade and consider it significant that countries such as Kenya, Zimbabwe and South Africa, all of which have been afflicted by drought, have been self-sufficient in food and usually have a surplus. They have in common a relatively well developed private sector producing consumer goods, coupled with good rural communications and well established trading centres.

I suggest that agricultural development projects should make provision also for training and for construction and loan capital for retail trade as well as for production. This would provide a local trading nucleus which would be an outlet for cash earnings and would serve to provide a local incentive within the project area, rather than in a remote town. It is fallacious to argue that such activity should be left to local entrepreneurs because, to a great extent, they lack the necessary capital or experience.

While this solution might conflict with some nationally cherished ideas of interdependence within the community and independence from external investment, I feel that it is only by adopting more liberal economic policies and specifically by encouraging the local production of consumer goods and the development of a rural distribution system, that the problem can be overcome and rural populations be brought fully into the economy.

D.H. ECCLES

56 St Margaret's Road,
Lowestoft NR32 4HT, UK

1. Perry, B. *Nature* **319**, 183 (1986).
2. Krebs, J.R. & Coe, M.J. *Nature* **317**, 13–14 (1985).
3. Sinclair, A.R.E. & Fryxell, J.M. *Can. J. Zool.* **63**, 987–994 (1985).
4. *Malawi Statistical Yearbook*, 1977 (Government Printer, Zomba).
5. Eccles, D.H. *Fishing News International*, April 1985, 46–47.

South Africa

SIR—Bender *et al.* (*Nature* **319**, 532; 1986) support the banning of South Africans from the World Archaeological Congress at Southampton and criticize your leading articles on the subject. Although they admit the existence of other repressive governments, they believe that South Africa should be singled out for two

reasons. The first is the view of men who think more of words than of what is happening in the real world. The Gulag is not institutionalized, nor are the death squads that operate in Latin American countries, nor are the mass slaughters that take place constantly in the Middle East. It follows plainly, at least to your correspondents, that archaeologists from these countries should be welcomed at Southampton. The logic of the argument escapes us.

The second reason given for excluding South African academics also involves some arcane dialectics. There are, say Bender *et al.*, “historical moments” when repressive governments become unstable and when outside pressure, such as an academic boycott, “may have some effect”, presumably in reforming them or destroying them. It is patently true that repressive governments are subject to “historical moments” when they become highly unstable; many of the dictatorships of South America are in precisely this state now, as are Uganda and several other countries in Africa, and half of the Middle East. Also, we all remember when Soviet tanks were required to preserve the governments of Hungary and of Czechoslovakia. As for Poland, its “historical moment” seems to be continuous. Bender *et al.* are interestingly unforthcoming about banning participants from any of these countries. On the contrary, as representatives of developing countries and of Eastern Europe they will be especially welcome. Again, we fail to follow the reasoning.

At the end of the letter we are asked to state “. . . what it is that you are asking us to tolerate”. The answer is simple: the presence at an international conference of about two dozen archaeologists most of whom oppose apartheid. What is more, they do so in South Africa where — unlike University College London — it takes courage.

J. MANDELSTAM

Microbiology Unit,

G.A. HARRISON

Department of Biological Anthropology,

E.T. HALL

Research Laboratory of Archaeology,
The University,
Oxford, UK

Too colourful

SIR—Vera Rich's report on the Chinese response to hydroengineering works on the Yangtse River (*Nature* **319**, 13; 1986) became a bit too colourful. The river with the gorges should not be equated with the Huang He (Yellow River).

STEPHEN H. EMBURY

Hematology Service,
Building 100, Room 263,
San Francisco General Hospital,
San Francisco,
California 94110, USA

CORRESPONDENCE

AIDS

SIR—In his letter on the origin of acquired immune deficiency syndrome (AIDS) (*Nature* 318, 100; 1985) Joseph Rosenior says that Japan was “virtually cut off economically and politically from the rest of the world until the latter part of the eighteenth century. There was hardly any commercial trade between Japan and Africa before the twentieth century.”

This elimination of other possible sources is precisely why the epidemiological studies of HTLV carriers, and the associated incidence of adult T-cell leukaemia/lymphoma (ATLL), in south-west Japan, are so strongly supportive of the hypothesis that the virus was introduced to Nagasaki by the Portuguese missionaries and traders in the 16th century; in particular, by the black Africans who came with them, and had frequent contact with local prostitutes. The historical roots of the still strong correlation ($r = 0.92$) between the incidence of ATLL and the density of Catholics are described in a letter by Hino *et al.*¹, with independent evidence clearly supportive of the suggestion by Gallo *et al.*² which Mr Rosenior criticizes.

MARK CANTLEY

Concentration Unit for
Biotechnology in Europe,
Commission of the
European Communities,
Rue de la Loi 200,
B-1049 Brussels, Belgium

1. Hino, S. *et al. Lancet* ii, 572 (1984).
2. Gallo, R.C. *et al. Lancet* ii, 962–963 (1983).

SIR—The letter from Fox and Cottler-Fox (*Nature* 319, 8; 1986) is scare-mongering although I would not deny that acquired immune deficiency syndrome (AIDS) is a worrisome new disease, and that as physicians we are relatively powerless against it at present. This has raised much speculation and controversy, some of which has been published in your columns.

Recently, it has been advised that donors of corneas to be used for transplantation in the United Kingdom should be screened for antibodies for LAV/ITLV III (lymphadenopathy-associated virus/human T-lymphotropic virus type III) and that material from those with antibodies should not be used. This advice has been widely heeded¹. Thus the supply of human corneas has been suddenly reduced, although the extent has yet to become apparent.

Although diseases such as rabies or Creutzfeldt–Jacob² syndrome can be transferred by corneal transplantation, a recent comprehensive review of the clinical manifestations of HTLV-III infection mentioned no evidence that AIDS can be transmitted in this way³.

It will thus be seen that AIDS has been added to the list of diseases where scare-

mongering has resulted in harm to patient care. This is always regrettable and reprehensible.

I do not entirely exclude the possibility that corneal transplantation might be a means of transmitting HTLV-III, merely that no such data yet exist. Furthermore, I do not believe that any scientist should fear being proved wrong by subsequent events. On the other hand, I do believe very strongly that it is wrong to make pronouncements or take actions which affect patients on the basis of no relevant data.

CHARLES CLAOUÉ

University of Bristol,
Department of Ophthalmology,
Bristol Eye Hospital,
Lower Maudlin Street,
Bristol B51 2LX, UK

1. Ritten, S. *Br. med. J.* 292, 62 (1986).
2. Kanski, J. *Clinical Ophthalmology*, 5.26, (Butterworths, London, 1984).
3. Melbye, M. *Br. med. J.* 292, 5–12 (1986).

Is a test ban unrealistic?

SIR—I note with concern your bald statement (*Nature* 319, 164; 1986) that “...aiming for a test ban now is...unrealistic. That task is for 1988...”

Why so? Surely there could be no more expeditious time than the present for such a critical move. A test ban is the most logical step to balanced multilateral nuclear disarmament, something that is supposed to appeal to all. Those who profess that the technical obstacles to confident verification are political, not technical, now greatly outnumber those who insist otherwise (at least outside the weapons laboratories and the Thatcher and Reagan administrations). Furthermore, we have a Soviet administration desperate for a reciprocal step to match its extended testing moratorium and its offer of on-site inspections of test sites and laboratories. Even if we adopt a cynical stance, the Soviet Union is highly vulnerable to having its bluff called. Either way, we would be foolish to sit back, prevaricate, and let the opportunities pass as our leaders seem to be intent on doing.

There is another reason. A Comprehensive Test-Ban Treaty would stop nuclear weapons spreading sideways, as well as upwards. In signing the Non-Proliferation Treaty, some 80 per cent of the United Nations membership agreed not to develop nuclear weapons *provided* the United States, the United Kingdom and the Soviet Union undertook to cease the arms race. Their disillusion is now tangible: all we three have done these past 15 years is escalate. Now it seems that the Soviets, at least, have lost their appetite. Even if they are ahead of us in megatonnage, as is commonly argued by those in favour of going ahead with development of the new generation of Western missiles,

the overkill capability of both sides renders the argument irrelevant. No, too much money is involved in the weapons industry, too many clever people in the weapons laboratories are intent on retaining their empires, too much inertia is built into the politicians' ability to regulate the race even if they had the will to do so. And too many scientists lose their intellectual honesty when lucrative defence contracts are proffered for fantastic schemes. We must break the chain. The test ban would be the way to start doing it. 1988 may be too late.

J.K. LEGGETT

122 Sinclair Road,
London W14 0NL, UK

Metric system

SIR—R.J. Bird's criticism (*Nature* 319, 8; 1986) of the metric system on the ground that it is more likely than the Imperial system to lead schoolchildren to mathematical ambiguities is invalid. His argument, in fact, has nothing to do with the metric system. 16 cm² and (16 cm)² can certainly be ambiguous to a young mind, but the same confusion arises when one writes in² for the square inch, which is not uncommon.

British schoolchildren can be told to use s.c. instead of cm² just like the more familiar c.c. for cm³ until they come to realize the logical superiority of the notation cm². The use of s.c. can be recommended outside elementary schools as well if need be.

KATUHISA SUZUKI

Department of Physics,
Osaka University,
Toyonaka 560, Japan

SIR—It is a rather strange experience to watch the defence for the non-metric traditions of the Anglo-American culture. The astonishing thing is not that some people cling to tradition, but that they sincerely believe that the metric units are almost meaningless when it comes to describing the physical world in immediately understandable terms. I can assure you that to a person who lives in a country that has been metric for generations, it is the system used in the United States and Britain that seems ridiculous—especially the Fahrenheit system so hotly defended by Mark W. Steele (*Nature* 318, 596; 1985). The obvious conclusion must be that the metric system serves as well for laymen when a transition period is over. In this country, acres, feet, gallons, pints and °F are as confusing to ordinary people as British thermal units are to physicists.

TRYGVE EKLUND

Protan A/S,
PO Box 420,
Tomtegt 36,
N-3001 Drammen,
Norway

Limits on mass of missing mass

Of the particles whose mass may account for the missing mass of the Universe, those known as axions are the most shadowy. What is known of stellar evolution surprisingly helps to define their properties.

IN the continuing if desultory search for the missing matter that might hold the Universe together gravitationally, the hypothetical (perhaps mythical) particle called the axion frequently crops up. The need for something along these lines follows from the evidence that clusters of galaxies appear to be more tightly bound to each other than the visible mass can account for, while the deceleration of the expansion of the Universe (poorly determined numerically) may also be more rapid than the visible mass allows. But the favourite constituent of the supposed dark matter of a few years ago, the neutrino with non-zero mass, is by no means as popular as it used to be.

For cosmologists, the attractiveness of the axion as the particle of which dark missing mass is made is much what that of neutrinos used to be. Again, the mass should be small compared with that of the electron, while the interaction of axions with other material particles should be very small, so that the chance that they would have been observed by accident in particle physics experiments would be negligible. One difficulty is that they should be produced prolifically in weak nuclear interactions, such as those responsible for beta-decay, in stellar interiors in particular. This, as it turns out, may be a difficulty. Or, more accurately, three physicists have now been able to show that what is known of the course of stellar evolution can be used to put limits on the properties of axions (see Dearborn, D.S.P., Schramm, D.N. and Steigman, G. *Phys. Rev. Lett.* **56**, 26; 1986).

Axions (if they exist) belong to a larger family of hypothetical particles called Goldstone bosons which owe their existence (if any) to the asymmetries in weak nuclear interactions, represented for example by the non-conservation of charge and parity combined. Such doubts as there may be about the reality of the particles are congruent with those applying to the theory of the weak interaction in general, which is more and more adequately confirmed as the weeks go by.

In some respects, axions stand in relation to photons as do muons in relation to electrons: they are more massive versions of the same thing. So axions are produced by the collision of energetic photons with electrons, as in the Compton effect, or with intact atoms, when they appear as the conversion products of the original photons. They may also appear instead of

photons in the collision of electrons with massive charged particles such as nucleons or nuclei, in a variant of the more familiar *bremstrahlung*. But there the analogy with photons stops; once produced, axions should not interact electromagnetically with other matter but only by the weak interaction, just like neutrinos. This explains why the chances of their having been detected as if by accident are very small.

This is where the cosmology comes in. Axions produced in very large numbers in the interiors of stars and wherever else energetic photons or electrons interact with other matter will, once created, rarely be destroyed again. So the Universe will contain a sea of axions, possibly concentrated (if they have a non-zero mass) where visible matter is also concentrated, near clusters of galaxies for example. Whether their presence would account for the present estimates of the missing mass, let alone suffice to close the Universe gravitationally, is another question whose answer would depend on their number and their individual mass.

The interest in what Dearborn, Schramm and Steigman have now done is that they start from the production end of the process. If axions are copiously produced within stars, they will necessarily carry away energy and so help cool the interiors of the stars producing them. The authors do not claim that this line of argument is novel, but they do appear to have used much better models for the evolution of stars than have previously been available for this purpose.

The stage in the evolution of a star at which the influence of the axions might be critical is that of the transition of a star from the long spell (as long as 10^{10} years for a star like the Sun) when hydrogen is the nuclear fuel to the succeeding stage, when the hydrogen is substantially exhausted and helium becomes the next most accessible nuclear fuel. The transition is recognized as that from a main-sequence star to a red giant.

Axion production could be important at this stage because helium burning requires a higher temperature at the centre of a star. As the hydrogen is depleted, the helium produced by hydrogen burning collapses under its own weight, with the result that both the temperature and the density at the centre of the star increase, as do the radius and the luminosity of the enlarged disk. Somewhere along the line, the central temperature is enough to ignite

the helium. But if the flux of axions is too great, the point may never be reached at which the helium is ignited; such an object would become redder, less luminous and less of a giant until it disappeared.

The observation that helium does ignite in stars no more massive than the Sun therefore provides a criterion for fixing an upper limit on axion production, at least with the assumption that axions once produced would travel through the envelope of a star as massive as the Sun essentially without interruption. Dearborn and his colleagues say that the fate of a star at this crucial stage will be delicately sensitive to the energy carried away by axions. Because the later stages of stellar evolution and those in which heavy elements are produced require that the transition from hydrogen to helium burning should be successfully accomplished, the character of the Universe as observed requires that there should not be too many axions.

The calculations, which entail the prediction of the evolution of a star beginning with the cosmic abundance of hydrogen and deuterium and nothing else for various assumptions about the properties of axions, are most of all a testimony to the sophistication of models of stellar evolution now in use. (The authors point out that, even so, no account is taken of the effects of rotation and of magnetic fields.) The inevitable outcome is that there is no unique upper limit for the mass of an axion, or of the other Goldstone bosons that might be substituted for it. But plausible assumptions about the stars involved and the constants that determine the production rates of axions give upper limits for the mass of an axion in the region of one electron-volt or less. Axions, if they exist at all, or other bosonic particles produced in the same way, are unlikely to carry mass greater than one millionth of the electron mass.

On the face of things, this should satisfy the cosmologists, who are looking for a particle whose interaction with other matter must be extremely feeble, which as a consequence must be long-lived and therefore abundant in the Universe, and which cannot carry more than a minute amount of mass without implying an uncomfortably large amount of missing dark matter. It will be interesting to see whether these new calculations suffice to rule out (or in) others of the notions that have recently been canvassed.

John Maddox

Astronomy

Hollow meteoroid streams

from David W. Hughes

IN December every year, meteors issue from the constellation of Gemini as Earth passes through the Geminid meteoroid stream. Radio meteoroids can be detected for about a week and the visual and photographic activity lasts slightly less long. Recent observations of the Geminid stream demonstrate that the conventional view of old meteor streams needs reconsidering. Instead of simply spreading as they age, the streams become hollow, literally forming pipes.

The Geminid meteor shower varies little from year to year, indicating that the stream is more than 500 years old. Recent observations with the Infrared Astronomical Satellite led to the detection of Asteroid 3200, which has a very similar orbit to that of the Geminid stream. It is still a matter of speculation as to whether Asteroid 3200 is the degassed nucleus core of the stream's parent comet or whether the Geminid comet choked to death by building up too thick an insulating dust crust and now appears asteroid-like.

Streams are formed from dust particles which have been ejected from the parent cometary nucleus with a spread of velocities, and thus slowly gain on or fall behind the comet until a complete annulus of dust eventually surrounds the cometary orbit. The orbits of these dust particles are perturbed by planetary gravitational fields, radiation pressure, the Poynting–Robertson effect and interactions between their electric charges and the interplanetary magnetic field. The Geminid stream makes no close approaches to any of the major planets, so the effect of perturbation is relatively straightforward. Jupiter is the major perturber with Saturn, Earth and Venus providing 4.1, 2.6 and 2.3 per cent, respectively, of the jovian effect.

J. Jones started a perturbation analysis of the Geminid stream by considering only Jupiter (*Mon. Not. R. astr. Soc.* **217**, 523; 1985). He replaced the stream by 71 meteoroids at random mean anomalies around the orbit, and then perturbed these particles by a point-mass Jupiter for a period of 5,000 years. He was expecting to find that the stream broadens with time with essentially a gaussian cross-section, but his results were “the causes of great astonishment”. Instead of Gaussians, he found that the cross-sections are like those of a thick-walled, somewhat squashed, pipe which is flattened parallel to the plane of the ecliptic. The in-ecliptic width of the pipe increases with time (at 1 AU from the Sun it is 0.006 AU across after 1,000 years and 0.022 AU across after

5,000 years, the other dimension reaching a minimum after 3,000 years but opening up again later). Adding the other planets to the calculation increases the cross-section of the pipe by a few per cent.

The Earth moves at about 0.017 AU per day around its orbit, so would take just over a day to cross a 5,000-year-old stream. If the Earth passes through a hollow stream, two brief periods of meteor activity would be observed, but as there is no reason why it should pass through the centre of the stream, the two bursts of activity would on average (assuming an elliptical cross-section) be separated by 64 per cent of the maximum possible time interval.

Has the predicted twin-peaked activity been observed for the Geminid stream? Radar data usually consist of many observations of faint meteors from a single station, whereas visual observations yield much lower fluxes but rely on many observers spaced over the globe. The addition of two, theoretical, separate gaussian activity curves convinced Jones that offsets of less than two days (corresponding to 16,000 years of perturbation) cannot be detected in radar fluxes from a single station. Combining results from non-identical, spaced stations and from year to year (which include intrinsic variations of activity) was thought to be too risky.

As the visual meteor shower only lasts half as long as the radar one the twin-peak nature of the proposed flux should be more obvious. A detailed study of visual Geminid data between 1970 and 1979 (see G.H. Spalding, *J. Br. astr. Ass.* **92**, 5; 1982) shows just such a bifurcation, a secondary maximum being clearly visible 0.8 days (0.014 AU) after the main maximum. This observation provides evidence that the Geminid stream is about 3,000 years old.

Jones stresses two other points. In the early history of the development of a meteor stream, other processes could mask the influence of planetary gravitational perturbation, but as the spreading caused by this effect is directly proportional to the age of the stream it will eventually dominate. If the stream has its aphelion near the orbit of Jupiter the perturbations causing hollowing are even more pronounced. Four streams immediately spring to mind: the Taurids; the Delta Aquarids; the Piscids; and the Quadrantids. The first three have double maxima that have previously been designated as the northern and southern branches of the showers. Jones thinks the picture is much simpler; one branch is seen when the Earth enters the thick pipe of meteoroid dust, the other when it leaves. The Quadrantid shower is extremely narrow (a characteristic that indicates youthfulness) and is therefore too young to have become hollow. □

David W. Hughes is Senior Lecturer in Astronomy and Physics at the University of Sheffield, Sheffield S3 7RH, UK.

Ecological modelling

Invasions of natural communities

from James A. Drake and Mark Williamson

INVADING species can affect natural communities and ecosystems in various ways. For example, the gypsy moth (*Porthetria dispar*) was introduced to North America in 1869; less than 100 years later it had defoliated more than 600,000 hectares of forest, and current predictions of damage to forest ecosystems in the northeastern United States are as high as 85 million hectares¹. There are many other examples of the disruptive potential of non-native species.

The international working group on modelling species invasions of the scientific committee on problems of the environment (SCOPE-ICSU) met recently* to discuss and to synthesize approaches to modelling the invasion process. The group is part of a programme which is examining

the ecological aspects of invasive species in natural and semi-natural ecosystems. It is to be hoped that accurate modelling of the invasion processes will lead to the development of cost-effective control measures; an understanding of the mechanisms governing species invasions also is essential if genetically engineered species are soon to be released into the environment².

Species invasion involves at least three distinct stages: arrival, establishment and dispersal. Elucidation of arrival and establishment mechanisms is the necessary first step in the modelling process; both deterministic and stochastic models are likely to be valuable (Stuart Pimm, University of Tennessee). Species movement throughout the biosphere is strongly influenced by patterns of human travel, and many more species are now candidates for long-distance dispersal than in the past. Pat-

*SCOPE working group on the modelling of the ecology of biological invasions. Fontana, North Carolina, 9–13 September 1985.

terms of human travel have been useful in modelling the spread of pathogens and can also help to determine invasion potentials. The probability of establishment is affected by the minimum viable population size and by community and environmental interactions. The population size required for establishment is influenced by thresholds such as the minimum number of individuals needed to sustain a mating system; genetic factors such as founder effects; life-history considerations such as age distribution, patterns of fecundity and factors governing recruitment; and the probability of extinction caused by stochastic fluctuations at low population levels.

After a species has established a viable population the dynamics of range expansion into adjacent habitats and across communities become important. Simon Levin (Cornell University) and Nils Stenseth (University of Oslo) identified key aspects of the dispersal stage, such as the stage of life-history where dispersal becomes important and rate of spread as a function of dispersal 'steps'. Stenseth pointed out that many surveys of species characteristics fail to show a strong correlation between species characteristics and colonization ability.

Models of species invasion generally fall into one of two broad categories with varying degrees of overlap. Levin described the geographical spread of a population in terms of a simple diffusion process³⁻⁵, where range expansion is a function of the irregular movement of individuals⁶, resulting in expanding (symmetrical or asymmetrical) concentric distributions (for example, the spread of muskrats in Europe⁷).

Invasion models should eventually include several parameters including dispersal mechanisms, temporal fluctuations, spatial heterogeneity and positive density dependence (Immanuel Noy-Meir, Hebrew University). Furthermore, theoretical representations of invasion must incorporate aspects of species biology, but only after a careful evaluation of which properties are important. Case studies of insect and mammal spread suggest some general patterns of range expansion beyond those predicted by simple diffusion models that are likely to be useful in formulating models based on Noy-Meir's classification (David Andow, University of Minnesota).

The second type of model considers community-level interactions once the invading species has arrived. A model of community assembly where community structure is generated by species invasions was presented by James Drake (Stanford University) in which communities that are initially vulnerable to colonizations and extinctions eventually become invasion-resistant. Characteristics of vulnerable and invulnerable communities may be

useful in predicting the fate of colonizing species. Accordingly, Pimm showed that in a community context species invasions and species removals tend to have substantially different effects. An interface between models of population spread and models of community dynamics is needed to predict the outcome of an invasion.

The only satisfactory models of population spread yet developed contain two parameters, the diffusion constant and the intrinsic rate of increase. It was generally agreed that models for management of ecosystems should have as few parameters as possible, and that they should be both measurable and interpretable. The parameters needed to predict success in establishment either vary from case to case⁸, or have yet to be identified.

The difficulty in constructing a list of attributes of successful invaders and predicting colonization success may stem from the need to consider such attributes in concert with community structure. Questions such as why some communities can be invaded whereas others cannot, and why some species can colonize a com-

munity at one time and not at another may not be answered until both species and community properties are incorporated. Multivariate analysis can show community interactions⁹, although surveys suggest that such interactions are rare¹⁰. □

1. Gerardi, M.H. & Grimm, J.K. *The History, Biology, Damage and Control of the Gypsy Moth* (Associated University Press, London, 1979).
2. Kolata, G. *Science* **229**, 34 (1985).
3. Kostitzin, V.A. *Mathematical Biology* (Harrap, London, 1939).
4. Skellam, J.G. *Biometrika* **38**, 196 (1951).
5. Murray, J.D. *Lectures on Nonlinear-Differential-Equation Models in Biology* (Clarendon, Oxford, 1977).
6. Okubo, A. *Diffusion and Ecological Problems: Mathematical Models* (Springer, Berlin 1980).
7. Elton, C.S. *The Ecology of Invasions by Animals and Plants* (Chapman and Hall, London, 1958).
8. Brown, A.H.D. & Marshall, D.R. *Proc. 2nd Int. Congr. Syst. Evol. Biol.* 351 (1981).
9. Williamson, M. *Symp. Br. Ecol. Soc.* (in the press).
10. Simberloff, D. in *Bioplectic Crisis in Ecological and Evolutionary Time*. (ed. Nitecki, M.H.) 53 (Academic, London, 1981).

James A. Drake is in the Department of Biological Sciences, University of California, Stanford, California 94305-2493, USA, working with Harold Mooney, chairman of the SCOPE programme; Mark Williamson is Professor of Biology at the University of York, York YO1 5DD, UK.

Anthropology

Tools, teeth and trampling

from Andrew Hill

ALTHOUGH almost trite, it is perhaps appropriate that those prehistorians who tease out the truth about past events and behaviour from scattered clues inadvertently left by our palaeolithic ancestors have been likened to famous fictional detectives. Recently, aspects of their work have taken an increasingly forensic slant, involving the detailed examination of stone instruments and of bones which may show evidence of injury caused by them. The report on page 768 of this issue by Behrensmeyer, Gordon and Yanagi¹ throws a shadow of doubt on what had been thought of as a breakthrough.

The basic problem is this: the earliest archaeological sites, classically those at Olduvai in Tanzania dated at almost 2 million years, consist of patterned concentrations of both stone artefacts and animal bones. Traditionally the association between them has been assumed to be functional — the bones were food remains, the artefacts were used to process them — with little question. This is not a trivial matter, for the Olduvai sites may constitute the earliest records of collective human activity. In addition, these assumptions have served as the basis for the elaboration of behavioural and social hypotheses concerning, for example, the origins of hunting and sharing. One manifestation of the greater rigour the subject has acquired in the past few years has been the criticism of such simple inferences.

Some have almost gone so far as to deny that these sites have anything to do with humans, the artefacts and bones having been brought together by such non-human factors as water action or carnivores. Others, although believing the association provides relevant data, see the need for independent evidence.

The breakthrough came from the study of scratches on fossil bone from Olduvai by Potts and Shipman² and by Bunn³ who also examined sites of a similar age at Koobi Fora in northern Kenya. The studies provided evidence that the bones had been cut by tools, possibly by butchery. Potts and Shipman² showed that the morphology of grooves experimentally cut by a stone edge appear very distinctive² at the high magnifications obtainable using a scanning electron microscope. The grooves contain fine internal striae that allow them to be readily distinguished from marks produced by carnivore teeth and other agents. This seemed to be the first direct evidence for a functional relation between the bones and the tools. The incidence of cutmarks was low and some overprinting of them by carnivore teeth marks suggested multiple origins for the Olduvai assemblages, or at least some non-human interference with them.

Further work has extended the range of this technique. The famous Acheulian sites of Torralba and Ambrona in Spain have long been thought unequivocally to

demonstrate the butchery of many animals. But when examined using the scanning electron microscope, only a very few bones at the sites show direct evidence of hominid activities, suggesting that humans did not play such a dominant part in the formation of the assemblages⁴. Another example is the clear cutmark evidence for the use of bone tools in North American mastodon butchery more than 10,000 years ago^{5,6}, which is important in the debate concerning the extinction of North American megafauna and in discussions of the reality and extent of Pleistocene bone technology. This, in turn, impinges on the controversy over the time of arrival of people in the New World.

Just as we were beginning to have real confidence in the results provided by the microscope techniques, Behrensmeyer and colleagues¹ examined the effects of trampling, a phenomenon likely to have been common by the lake shores on which these early sites accumulated as herds of beasts came down to drink. Fiorillo⁷ and Oliver⁸ have recently commented on trampling as a possible cause of pseudo-cutmarks. Behrensmeyer *et al.* report the results of controlled trampling experiments. First, abrasion from sand grains causes grooves on the bones that microscopically mimic those made by slicing with a stone tool. Trampling also removes distinctive features of pre-existing true cutmarks, which may explain the relatively low incidence of identifiable cutmarks in these early assemblages. Behrensmeyer *et al.* conclude that microscopic features alone are insufficient evidence for stone-

tool use.

The authors discuss other criteria that may assist in the problem. Human activities could plausibly produce cuts in characteristic locations on bones, but trampling also affects these same areas differentially, as a consequence of bone shape. Trampling can be eliminated when sand and gravel are absent from the environment of burial, but I am not sure that such a context exists by African lakes. Also, in the dry season, fine-grained sediments are baked hard, forming a firm substrate against which bones may be trampled under sharp hooves. Particular patterns of multiple marks seem not to be characteristic of trampling; therefore, when they occur, human involvement can be cautiously assumed.

This is not much to be left with, and the new work clearly raises serious problems. It may also raise criticism from other workers in the field, which could lead to a clarification and refinement of demonstrable criteria for distinguishing between purposive cuts produced by tools in human hands and incidental cuts made by fortuitous feet. □

1. Behrensmeyer, A.K., Gordon, K. & Yanagi, G. *Nature* **319**, 768 (1986).
2. Potts, R. & Shipman, P. *Nature* **291**, 577 (1981).
3. Bunn, H. *Nature* **291**, 574 (1981).
4. Shipman, P. & Rose, J.J. *J. Archaeol. Sci.* **10**, 465, (1983).
5. Fisher, D.C. *Nature* **308**, 271 (1984).
6. Shipman, P., Fisher, D.C. & Rose, J.J. *Paleobiology* **10**, 358 (1984).
7. Fiorillo, A.R. *Current Res.* **1**, (1984).
8. Oliver, J.S. *Int. Conf. Bone Modif. Abstr.* 27 (1983).

Andrew Hill is in the Department of Anthropology, Yale University, PO Box 2114 Yale Station, New Haven, Connecticut 06520, USA.

Earthquake modelling

Caustics in seismology

from Francis Wright

IF it were possible to take a seismic photograph in a region of strong earthquake activity, it might look something like Fig. 1. Or rather, that is what one coherent frequency component might look like. John Nye has made detailed experimental and theoretical studies of the propagation of light through small water droplets^{1,2} (which is how he produced the photograph in Fig. 1) and has now, using catastrophe theory, applied some of that knowledge to the propagation of seismic waves through the Earth³.

Wave energy rarely propagates along

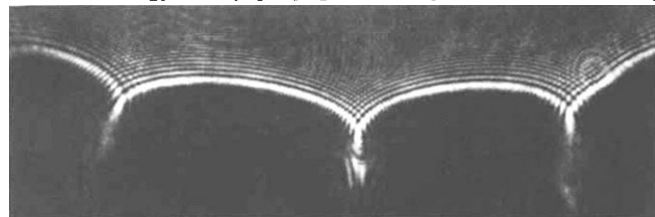


Fig. 1 A caustic showing cusps and diffraction (photograph, J. Nye).

straight lines. Variations in the medium cause rays to curve, which in turn can cause neighbouring rays to converge or focus to produce high-intensity regions or caustics. An example of this process is the focusing of light rays caused by varying air temperature, which is responsible for mirages. The very bright curve dividing the brighter top half from the darker bottom half of Fig. 1 is a caustic. The fine structure or diffraction pattern is produced by the 'wavelets' travelling along neighbouring rays interfering with each other. The scale of the diffraction pattern is of the order of the wavelength — microns for light but kilometres for earthquakes.

Nye analyses seismic caustics by starting out with a spherically symmetrical Earth and then con-

sidering the effect of irregularities. It is quite easy to construct caustics in highly symmetrical models, but on breaking the symmetry there are myriad possibilities for the metamorphoses of the caustics. Catastrophe theory^{4,5}, in the guise of 'catastrophe optics'⁶, can be applied to show what forms the caustic is likely to have within any small region. There are very few such forms: two in two dimensions (on the surface of the Earth); five in three dimensions (inside the Earth); seven in four dimensions (inside a family of Earths); and so on.

Nye gives two examples of seismic caustic systems (Fig. 2 *a* and *b*). One of these systems (*a*) is caused by sudden increases in seismic velocity with depth in the transition zone of the Earth's mantle. In the

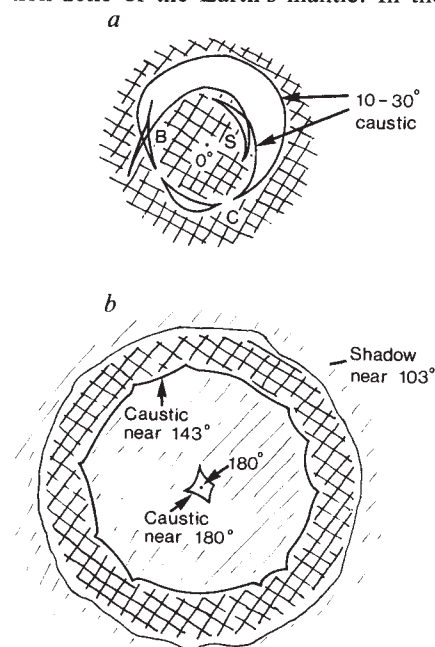


Fig. 2 Schematic seismic caustics viewed *a*, from directly above and *b*, below (opposite) the earthquake source, roughly to scale (based on ref. 3). Shading is lighter for areas of higher seismic intensity.

spherical model this results in two concentric caustic circles on the Earth's surface of angular radius between 10° and 30°. Each of these circles is the line on which a caustic surface meets the Earth's surface, and these two caustic surfaces meet inside the Earth along a cusped edge.

What happens to this caustic when the symmetry is relaxed? Locally, probably nothing significant, as the cusped caustic surface corresponds to a stable catastrophe. But globally, the caustic will cease to be circular and the cusped edge may pierce the surface, making the two caustic lines on the surface meet in a cusp (such as C in Fig. 2*a*).

If the distortion of the caustic is sufficient, the cusped edge inside the Earth may fold back on itself to produce a local caustic configuration called a swallowtail, which also corresponds to a stable catastrophe.

trophe in three dimensions. This conformation of the caustic surface would explain a caustic line on the surface that doubles back on itself twice at two cusps (S in Fig. 2a). In fact, observations in the western United States suggest caustic lines doubling back on themselves three times (B in Fig. 2a). This could be explained in catastrophe terms by invoking the stable existence in a family of Earth interiors (a four-dimensional space) of a 'butterfly' catastrophe that occurs near to our real Earth, although Nye only hints at this possibility.

The other of Nye's examples of a caustic system (Fig. 2b) occurs on the other side of the Earth from the source, and is the result of refraction occurring twice as rays pass through the core. This is somewhat analogous to that most familiar caustic, the rainbow, which is caused by refraction of light passing into and out of raindrops, but with an internal reflection in between (see refs 7 and 8). In a spherical model Earth, the core casts a circular shadow with an angular radius of 103° , and causes focusing onto a circular 'rainbow' caustic at 143° and a caustic point at 180° . The

caustic point is analogous to the bright spot at the centre of the shadow of a circular object. It is highly unstable, and on relaxing the symmetry it will typically 'explode' into a loop with four outward-pointing cusps, which is probably about 10° across on the Earth.

By contrast, the caustic at 143° is locally stable, although when the symmetry is relaxed it may develop cusps, and look locally something like Fig. 1. Nye argues that bumps on the core-mantle boundary could produce cusps protruding by up to 2° from the unperturbed caustic, although in observations these cusps are then smeared out by averaging over seismic events. □

1. Nye, J.F. *Proc. R. Soc. A* **361**, 21 (1978).
2. Nye, J.F. *Phil. Trans. R. Soc. A* **292**, 25 (1979).
3. Nye, J.F. *Geophys. J. R. astr. Soc.* **83**, 477 (1985).
4. Thom, R. *Structural Stability and Morphogenesis: an Outline of a General Theory of Models* (Benjamin, Reading, Massachusetts, 1975).
5. Poston, T. & Stewart, I. N. *Catastrophe Theory and its Applications* (Pitman, London, 1978).
6. Berry, M.V. & Upstill, C. *Prog. Optics* **18**, 257 (1980).
7. Marston, P.L. & Trinh, E.H. *Nature* **312**, 529 (1984).
8. Nye, J.F. *Nature* **312**, 531 (1984).

Francis Wright is in the School of Mathematical Sciences, Queen Mary College, University of London, Mile End Road, London E1 4NS, UK.

Learning

A step nearer a neural substrate

from Graham V. Goddard

ON PAGE 774 of this issue, Morris *et al.*¹ provide new and provocative evidence that learning can be accounted for in terms of known and specific changes in the synapses of the brain. They report the first *in vivo* experiment that suggests the phenomenon of long-term potentiation — the long-lasting increase in synaptic transmission first described by Bliss and Lømo² — is essential to the substrate of a learned behaviour.

Any physical substrate of learning needs to be durable and of sufficient flexibility to allow for the many possible variations in learned behaviour. Since the idea was first proposed in 1893 by Tanzi³, it has been commonly thought that the required durability and flexibility would be obtained if the power of particular neurones to excite certain others could be increased by structural changes at their points of communication, the synapses. In 1949 Hebb⁴ published a refined version of this hypothesis with more theoretical details, showing how selective changes at synapses might arise from experience and could account for quite complex behaviour. It was not until 1973, however, that Bliss and Lømo² demonstrated in an adult mammalian brain an appropriate activity-dependent long-lasting increase in the power of a specific set of synapses, which they called long-term potentiation. Up till now, this potentiation can be demonstrat-

ed only in an experimental paradigm that requires electrical brain stimulation and recording through appropriately placed electrodes. Although the changes may persist for weeks or months, a direct involvement in behaviour has not so far been established.

Other 'plasticities', both structural and functional, have been demonstrated in the nervous system but they may not be the same as that underlying learning. Either they involve responses to brain damage, demonstrate a critical role for experience in the differential development of an immature brain, are too short in duration, or are decremental changes better suited to explaining the reduced response to frequently occurring stimuli than learning in a choice situation. An example of such habituation processes is the gill withdrawal reflex in *Aplysia*, where long-lasting changes in synaptic efficacy have also been reported⁵ and the desired evidence at the synaptic level may soon be provided⁶.

To produce the changes in synaptic power that are seen in long-term potentiation requires the cooperative action of many converging excitatory synapses, with relatively little opposing inhibition. These conditions bear a formal resemblance to those proposed by Hebb⁴, and suggest that long-term potentiation might indeed serve as the substrate of associative

learning, in which a certain combination of stimulus conditions are assumed to be necessary for a particular learned alteration in behaviour to occur.

The evidence now presented by Morris *et al.*¹ is important because it indicates that long-term potentiation is indeed the substrate of a learned behaviour because the learning cannot occur in its absence. The logic of their experiment is based on the recent important discovery by Collingridge, Kehl and McLennan⁷ that long-term potentiation depends critically on the activation of *N*-methyl-D-aspartate (NMDA) receptors, one of several types of receptors responding to the neurotransmitter chemical glutamate. Normal excitation of neurones activated by glutamate released at presynaptic terminals is mediated by other types of glutamate receptors, the kainate and quisqualate types, and does not require activation of NMDA receptors, even after long-term potentiation. But at the moment when long-term potentiation is set up, activation of NMDA receptors is required.

Morris and his colleagues show that blocking NMDA receptors by topical administration of the selective antagonist aminophosphonovaleric acid (AP5) not only prevents long-term potentiation but also prevents learning in a behavioural situation. The blocker does not prevent normal synaptic transmission or normal behaviour. The effect on learning was demonstrated using swimming rats in two situations. In the first, the rats were required to learn the location of a submerged platform in a swimming pool from which there was no other exit. In the second, they had to learn to go to a secure platform which was distinguished by visual markers from an insecure one. The NMDA blocker prevented the rats from learning about the location of the platform but did not interfere with learning about the visual cues. This is significant because it shows that the failure of place learning is not due to some general disability but is a very selective and specific effect.

These results, however, raise an interesting theoretical difficulty: if all learning is based on long-term potentiation, why is all learning not prevented by the NMDA blockade? There are at least two possible answers to this problem. First, the substrate of visual-cue learning may have a completely different neurochemical

Erratum

The last sentence of the first paragraph of A.J. Dessler's article "Does Uranus have a magnetic field?" (*Nature* 16 January, p.174) should have read "The present fascinating state of confusion is a result of earlier observations of Uranus from Earth orbit and some recent data from Voyager that do not confirm recent expectations." Also ref.1 should have included Durrance, S.T. & Moos, H.W. *Nature* **299**, 428; 1982 and ref.4 should have had G.R. Smith as a coauthor.

mechanism from that of place learning. Alternatively, and more plausibly, in the experiments which I have just described the NMDA blocker could have reached areas of the brain involved in place learning but not areas involved in visual-cue learning. It is known that, in the rat, an intact hippocampus is required for place learning and not for visual-cue learning⁸. Morris *et al.* blocked NMDA receptors by infusing AP5 into the lateral ventricle, which is very close to the hippocampus. It is possible that the areas of the brain containing synapses crucial for visual-cue learning are more remote from the lateral ventricle and so received less drug.

The next step will be a double dissociation experiment in which visual-cue learning is blocked by a treatment that leaves place learning unimpaired. If this can be done by administering the same

NMDA blocker by a different route, the evidence that long-term potentiation of synapses is the basis of behavioural learning will become strong indeed. □

1. Morris, R.G.M., Anderson, E., Lynch, G.S. & Baudry, M. *Nature* **319**, 774 (1986).
2. Bliss, T.V.P. & Lomo, T. *J. Physiol., Lond.* **332**, 331 (1986).
3. Tanzi, cited in Ramón y Cajal, S. *Histologie du Système Nerveux de l'Homme et des Vertébrés*. Vol. 2, p. 886 (Maloine, Paris, 1911).
4. Hebb, D.O. *The Organization of Behavior, a Neuropsychological Theory* (Wiley, New York, 1949).
5. Carew, T.J., Pinsky, H.M. & Kandel, E.R. *Science* **175**, 451 (1971).
6. Walters, E.T. & Byrne, J.H. *J. Neurosci.* **5**, 662 (1985).
7. Collingridge, G.L., Kehl, S.J. & McLennan, H. *J. Physiol., Lond.* **334**, 19 (1983).
8. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Oxford University Press, 1978).

Graham V. Goddard is at the Department of Psychology and the Neuroscience Centre, University of Otago, PO Box 56, Dunedin, New Zealand.

Oncogenes

Tricks with tyrosine kinases

from J. Michael Bishop

PROTEIN kinases that phosphorylate tyrosine were first identified as products of retroviral oncogenes, but the provenance of these enzymes soon expanded greatly¹. The members of the family of protein-tyrosine kinases play vital roles in the regulation of cellular phenotype — some as cell-surface receptors for polypeptide growth factors, others having intracellular functions whose natures have yet to be described. For a while protein-tyrosine kinases seemed destined to fulfill the prediction that the first shall be last and the last first: biologists interested only in human carcinogenesis paid little attention to phosphorylation of tyrosine; point mutations in *c-ras* and rearrangements of *c-myc* were all the rage. But the protein-tyrosine kinases now return to the fore — not only because they brook so large in the economy of normal cells, but because they have at last been implicated in tumorigenesis of various origins. The latest example is reported elsewhere in this issue², and is in some ways the most beguiling yet encountered.

Barbacid and colleagues had detected and isolated an active oncogene (*onc-D*) from the DNA of a human carcinoma of the colon³, but its large size (more than 30 kilobase pairs) delayed its characterization. First attempts to elucidate the sequence of its nucleotides uncovered only a small swatch of the coding domain. But those of us who heard descriptions of that swatch at symposia went home with our heads spinning — the sequence was the same as part of that for a non-muscle tropomyosin. The uncertain implication was that a component of the cytoskeleton might elicit neoplastic growth: the un-

spoken truth was that few were willing to imagine how such a thing could happen.

Martin-Zanca *et al.*² now report the nucleotide sequence of a complementary DNA that contains the entire coding domain of *onc-D* and that can consequently transform cells in culture. The protein encoded by this DNA has a startling composition: the amino-terminal quarter is represented by the tropomyosin I have already mentioned, the remainder by the business end of a protein-tyrosine kinase. The kinase domain of the protein is branded as a transmembrane receptor by features of its structure, but it is not a receptor that is presently recognizable; whatever ligand-binding domain it normally possesses has been deleted by the fusion with tropomyosin. Nothing more can yet be said of how badly mangled the receptor protein is because its normal counterpart has not yet been characterized.

It is clear, then, that *onc-D* arises from recombination between separate genetic loci, perhaps from translocation between different chromosomes. This is an example of how molecular biologists can blunder on chromosomal rearrangements not yet sighted by cytogeneticists. In rare instances, the events of experimental gene transfer have themselves generated abnormal genes that are active in assays for oncogenes^{4,5}. Martin-Zanca *et al.* show that the rearrangement represented by *onc-D* can be found in a specimen of the tumour from which the gene apparently originated, but not in adjacent normal tissue. Therefore, the oncogene arose as a somatic lesion at some point during the genesis or growth of the tumour.

Why should the chimaeric protein encoded by *onc-D* transform cells? Most observers would place the blame on the protein kinase, but the tropomyosin component seems likely to be an accessory to the fact — by directing the hybrid protein to an improper location within the cell; by causing it to interact with illegitimate substrates; or by replacing a portion of the kinase molecule that would otherwise exercise allosteric control over enzymatic activity. These thoughts recall the structure of two retroviral oncogenes: *v-fgr*, a fusion between genes encoding the cytoskeletal protein γ -actin and a protein-tyrosine kinase (not a transmembrane receptor, in this instance)⁶; and *v-erb-B*, encoding a severely truncated version of the receptor for epidermal growth factor in which allosteric control may have been discarded by deletion of the ligand-binding domain⁷.

The isolate of *onc-D* described by Martin-Zanca *et al.* is for the moment one of a kind. But abnormalities of other protein-tyrosine kinases have also been uncovered in tumour cells. For example, the translocation (9;22) that gives rise to the Philadelphia chromosome (22q⁻) of human chronic myelogenous leukaemia creates a chimaeric protein, encoded in part by a newly recognized region of chromosome 22 (tentatively called *bcr*) and in part by the bulk of the proto-oncogene *c-abl* — including its protein-tyrosine kinase domain^{8,9}. The fusion is thought to unbridge or otherwise alter the enzymatic activity of the protein kinase¹⁰.

In a second example, insertional mutagenesis by retroviral DNA augments the production of a newly described protein-tyrosine kinase in the murine T-cell leukaemia lines LSTRA and Thy-19 (refs 11,12). The two groups that made these observations have regrettably given different names to the gene that encodes the kinase (*lsk⁺* and *ick*), but they agree that the exceptional abundance of the kinase may contribute to neoplastic growth of the leukaemia cells. And finally, gene transfer with DNA from tumour cells has turned up two new oncogenes that encode protein-tyrosine kinases: *neu*, isolated from chemically induced neuro- and glioblastomas of rats¹³; and *met*, found in human tumour cells that have suffered chemical mutagenesis following establishment in culture¹⁴. The *neu* gene encodes a cell-surface receptor for a presently unknown ligand^{15,16}. The version of *neu* isolated from tumour cells presumably bears a mutation that would account for the ability of the gene to transform cells in culture, but neither the nature nor the effect of the presumed mutation is known. Little is known of the protein encoded by *met* — so little, in fact, that the present claim to fame of the gene lies mainly in its potential use as a polymorphic marker for cystic fibrosis¹⁷.

The preceding description may seem something of a thicket to the general reader. But it carries a lesson. Much ink has been spilled over the issue of whether oncogenes are activated by events that change gene expression or mutations that change gene products. The ink has been wasted. Even in the microcosm of protein-tyrosine kinases described here, both routes of activation apparently occur.

There is yet another twist to the story of *onc-D*. It has become common to find more than one tropomyosin produced from a single gene by alternative splicing^{18,19}. Martin-Zanca *et al.* hint that the same is true of the tropomyosin gene in *onc-D*. If so, two possibilities arise. First, *onc-D* may be active only in cells that produce non-muscle tropomyosin: the pattern of splicing in muscle cells might exclude the protein-tyrosine kinase domain from the gene product. Second, genes expressed through alternative pathways of splicing may be prone to genetic rearrangements of the type that engender *onc-D* (although it is not clear how such a predisposition would operate). There is no obvious precedent for this suggestion, but it is nevertheless provocative that translocation has excluded alternatively spliced exons from the expression of proto-oncogenes in at least two instances: the four (or more) alternative first exons of *c-abl* in the Philadelphia chromosome²⁰; and the last (non-muscle) exon of the tropomyosin gene included in *onc-D*². The discoverers of *onc-D* suggest that the gene be rechristened *trk* and that the acronym be pronounced 'track'. But as a bemused bystander, I find 'trick' more apt. □

- Hunter, T. & Cooper, J.A. *A. Rev. Biochem.* **54**, 897 (1985).
- Martin-Zanca, D., Hughes, S.H. & Barbacid, M. *Nature* **319**, 743 (1986).
- Oulciani, S., Santos, E., Lauver, A.V., Long, L.K. & Barbacid, M. *Nature* **300**, 539 (1982).
- Santos, E., Reddy, E.P., Pulciani, S., Feldmann, R.J. & Barbacid, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4679 (1983).
- Takahashi, M., Ritz, J. & Cooper, G.M. *Cell* **41**, 581 (1985).
- Naharro, G., Robbins, K.C. & Reddy, E.P. *Science* **223**, 63 (1984).
- Ulrich, A. *et al. Nature* **309**, 418 (1984).
- Shivelman, E., Lifshitz, B., Gale, R.P. & Canaani, E. *Nature* **315**, 550 (1985).
- Stam, K. *et al. New Engl. J. Med.* **313**, 1429 (1985).
- Konopka, J.B., Watanabe, S.M., Singer, J.W., Collins, S.J. & Witte, O.N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1810 (1985).
- Marth, J.D., Peet, R., Krebs, E.G. & Perlmutter, R.M. *Cell* **43**, 393 (1985).
- Voronova, A.F. & Sefton, B.M. *Nature* (in the press).
- Schechter, A.L. *et al. Nature* **312**, 513 (1984).
- Dean, M. *et al. Nature* **318**, 385 (1985).
- Bargmann, C.I., Hung, M.-C. & Weinberg, R.A. *Nature* **319**, 226 (1986).
- Yamamoto, T. *et al. Nature* **319**, 230 (1986).
- White, R. *Nature* **318**, 382 (1985).
- Basi, G.S., Boardman, M. & Sorti, R. *Molec. cell. Biol.* **4**, 1818 (1984).
- Ruiz-Opazo, N., Weinberger, J. & Nadal-Ginard, B. *Nature* **315**, 67 (1985).
- Ben-Neriah, Y. *et al. Cell* (in the press).

J. Michael Bishop is at the Department of Microbiology and Immunology and the Hooper Research Foundation, University of California Medical Center, San Francisco, California 94143, USA.

Palaeobotany

On follicles and flowers

from Margaret Collinson

INNOVATIONS in reproductive biology, including floral organization and function, may account for much of the success of the angiosperms — the flowering plants. Our knowledge of the early diversification of these plants has relied heavily for many years on extrapolation from the fossil pollen record. Flowers themselves have been considered unlikely fossils because of their compound structure with delicate and easily shed organs. Certainly most known fossil flowers have been preserved in a manner that has precluded serious scientific study. Recent publication of material from Japan¹, Europe² and North America^{3,4} now redresses this balance and provides the first substantial evidence for early angiosperm flowers and their subsequent diversification.

Protection of ovules by an enclosing structure (the carpel) with a receptive surface (the stigma) permits control over pollination and hence fertilization. This, combined with co-evolution of faithful pollinating vectors and dispersal agencies, such as insects and birds, may be largely responsible for the diversity and the dominance of the flowering plants today. Studies of isolated pollen, reviewed by Muller⁵, suggest that these strategies first appeared at least 115 million years ago. It is therefore of interest to elucidate the nature of the earliest and most generalized members of the flowering plants and to follow their subsequent diversification which has resulted in the amazing variety that we see today.

Essential elements in one widely held hypothesis of primitive flowering plant morphology⁶ are many free, spirally arranged, conduplicate carpels (ovule-bearing structures resembling folded leaves) maturing into follicles (many-seeded fruits splitting along a single, adaxial suture). This hypothesis, based on neontological study, could be tested using a stratigraphic (time-based) sequence of elaboration in floral organization from the fossil record. It may seem surprising that such an obvious test has not yet been applied but there are probably two reasons for this. First, Darwin's reference to the origin of the angiosperms as an "abominable mystery" has directed palaeobotanical attention to the search for potential ancestors rather than to the study of early fossil angiosperms themselves; second, the fossil material needed for such an analysis is only now becoming available.

Nishida¹ recently reported a permineralized floral axis from the mid-Cretaceous (approximately 90 million years ago) of Japan. This particular preservation state

reveals details of the vascularization and cellular structure of many bitegmic ovules in many spirally arranged follicles, borne on a concave receptacle with a woody axis containing scalariform vessels with scalariform perforation plates.

Slightly older fossils (approximately 95 million years) have been described from the western interior of North America^{3,4}. One of these³, *Lesqueria*, is preserved as



Reconstruction of a leafy twig bearing a multifollicular axis of *Archaeanthus linnenbergeri* from the mid-Cretaceous of central Kansas, North America (from ref. 4).

three-dimensional moulds in sandstone, which are less informative than the Japanese specimen but do reveal a similar floral organization, although with a dome-shaped receptacle and many more follicles. As *Lesqueria* has previously been considered to be a bennetitalean fructification, careful reinvestigation of similar fossils is required.

The *Archaeanthus* plant⁴ (see figure) is, in contrast, known in considerable detail from what might seem to be rather poorly preserved compression/impression fossils. Many organs of this plant are found together in a narrow fossiliferous horizon, and scars on the floral axis are consistent with the attachment of dispersed leaves, bud scales and perianth. The reconstruction suggests that the flower possessed three sepals, six or nine petals, up to 60 stamens represented only by scars and more than 100 spirally arranged carpels (each containing many ovules) that matured into stalked follicles. The flower was large, the multifollicle up to 15 cm high and the petals up to 8 cm long.

This flower most closely resembles modern magnolid flowers, but this type of

structure conforms more closely to the hypothetical primitive form than do flowers of any modern member of the group. It is extremely exciting and significant to find such a form, predicted on the basis of neontological studies but absent from any modern plant, in this early phase of the angiosperm fossil record.

Several other flower forms are contemporary with these multifollicles, for example, a pentamerous form with free whorled sepals, petals and stamens, and loosely fused carpels⁷. These argue against a naive interpretation which might have considered the multifollicles as the most primitive flower form. Furthermore, deductions⁸ from a comparative morphological assessment of fossil and modern pollen suggest that ancestral complexes, with descendants in most major groups of modern dicotyledons, had already differentiated before the time when the early flower forms we now recognize existed.

Clearly, the multifollicular floral organization was present at a very early stage of angiosperm evolution in several areas of the globe. Our knowledge of the subsequent diversification of early flowers has been greatly elaborated by the exciting finds of small (2–3 mm) charcoaled flowers described by Friis² from the upper Cretaceous (approximately 78 million years ago) of Sweden. There are more than 20 forms, mostly regular and bisexual with whorled floral parts in definite numbers. None have free carpels and many have an inferior ovary. Some show specializations for wind pollination and are close relatives of the present-day Juglandales; others were probably insect pollinated, as they had a nectariferous disk, and are related to the Saxifragales.

Although fossils of Tertiary flowers have always been more common than Cretaceous forms, there has also been a con-

siderable recent expansion in our knowledge of them. Crepet⁹, working mainly with North American material, has demonstrated various elaborate floral forms such as zygomorphic flowers, brush flowers and long corolla tubes that are consistent with the presence of faithful insect pollinators. These were in existence by the middle Eocene (approximately 45 million years ago). About 70 flower forms from Messel, a West German Eocene site, are currently being investigated by Schaarschmidt¹⁰. These combined data are now sufficient to provide an initial overview of the elaboration over a considerable period of time in floral organization and the related pollination strategies (W.L. Crepet and E.M. Friis, personal communication).

A few years ago the response to questions concerning fossil flowers would have been at best pessimistic and at worst dismissive. This has suddenly changed with the description of fossils from the lower/upper Cretaceous boundary onwards that have revealed significant and fundamental information concerning early angiosperm flowers and their subsequent diversification. The present outlook is much more rosy. □

1. Nishida, H. *Nature* **318**, 58 (1985).
2. Friis, E.M. *Ann. Mo. bot. Gdn* **71**, 403 (1985).
3. Crane, P.R. & Dilcher, D.L. *Ann. Mo. bot. Gdn* **71**, 384 (1985).
4. Dilcher, D.L. & Crane, P.R. *Ann. Mo. bot. Gdn* **71**, 351 (1985).
5. Muller, J. *Ann. Mo. bot. Gdn* **71**, 419 (1985).
6. Takhtajan, A. *Flowering Plants, Origin and Dispersal* (Oliver & Boyd, Edinburgh, 1969).
7. Basinger, J.F. & Dilcher, D.L. *Science* **224**, 511 (1984).
8. Walker, J.W. & Walker, A.G. *Ann. Mo. bot. Gdn* **71**, 464 (1985).
9. Crepet, W.L. *Ann. Mo. bot. Gdn* **71**, 607 (1985).
10. Schaarschmidt, F. *Ann. Mo. bot. Gdn* **71**, 599 (1985).

Margaret Collinson is in the Department of Biology, King's College London (KQC), 68 Half Moon Lane, London SE24 9JF, UK.

Organelle transport

Moving in different directions

from Peter Hollenbeck

ALTHOUGH all animal cells can redistribute their contents by the rapid transport of membrane-bound organelles, two properties of neurones make them particularly useful for the study of this process. First, their sheer size and length make transport a primary occupation¹; second, their highly polarized structure makes clear the intrinsic polarity of organelle transport, which is disguised in most other cell types. Although it is known that different populations of organelles travel in the anterograde from the retrograde direction, the nature of the motor driving the movements remains uncertain. Are anterograde and retrograde transport supported by identical force-transduction

mechanisms or by different ones? Recent studies using a reconstituted system of movement suggest the latter possibility².

I recently described in these columns³ how, using an *in vitro* assay for microtubule-based movement, Vale and co-workers identified a force-generating protein, kinesin, which may play a part in organelle transport along the axons of nerves⁴. Having established that kinesin can cause movement along microtubules, these authors now approach the question of the polarity of organelle transport and provide evidence that different force-generating systems are responsible for transport in the anterograde and retrograde directions².

Microtubules, polarized filaments with 'plus' and 'minus' ends defined by the *in vitro* assembly properties of their tubulin subunits, form the tracks along which rapid organelle transport occurs. As virtually all the microtubules in the axon are oriented with their plus ends in the direction of the synapse, and organelle traffic can proceed both towards and away from the synapse, it follows that transport must occur in two different fashions with respect to the polarity of the microtubule substrate. But when assayed for its ability to translocate artificial organelles (in the form of polystyrene beads) along isolated microtubules *in vitro*, kinesin supports movement in only one direction².

Vale *et al.*² have now determined to which of the two directions this corresponds *in vivo*. Using isolated centrosomes (microtubule-organizing centres in the cell) as seeds for microtubule growth *in vitro*, it is possible to obtain radial arrays of microtubules with all the plus ends pointing outwards⁵. Using such polarized arrays as tracks for the movement of kinesin-coated beads, the authors show that this protein supports movement only towards the plus ends of the microtubules, which corresponds to movement from the neuronal cell body towards the synapse *in vivo*. Hence it is possible that kinesin is the motor for anterograde transport.

If beads are treated with a crude extract of axonal cytoplasm instead of with pure kinesin, their transport along microtubule arrays changes. Instead of travelling in one direction, they move towards both the plus and the minus ends of the microtubules. Furthermore, movement of beads in the two directions in the reconstituted system differs from the native preparation in several ways: first, the velocity of movement is greater in the retrograde than in the anterograde direction; second, pharmacological agents that inhibit retrograde movement leave anterograde movement unaffected; third, monoclonal antibodies directed against kinesin can be used to remove the anterograde translocation activity from crude extracts while leaving the retrograde activity unchanged. This indicates that the cell contains separate translocation factors for movement towards the plus and minus ends of microtubules, and thus suggests that different mechanochemical motors mediate transport in the two directions *in vivo*.

The possibility that this might be the case has been suggested only indirectly by physiological studies, and would not be predicted by ultrastructural studies in which the crossbridges between the organelles and the microtubules appear identical regardless of the direction in which the organelles are moving at the time of fixation⁶.

A final point of interest raised by Vale *et al.*² is the difference in motile behaviour between artificial and endogenous organ-

elles when crude axoplasmic extracts are applied to polar microtubule arrays. Natural organelles exhibit smooth continuous movement in only one direction, whereas polystyrene beads have a jerky motion, reversing direction frequently. The beads, which have a negative surface charge, might be expected nonspecifically to bind many proteins from the crude extract. Organelle movement in this assay is consistent with the presence on the surface of force-generating machinery for both directions of movement. This raises the possibility that the surface properties

of endogenous organelles specify which direction they will travel by determining which mechanochemical system they can bind, and suggests a simple mechanism by which the cell can direct traffic. □

1. Brady, S.T., Lasek, R.J. & Allen, R.D. *Cell Motil.* **5**, 81 (1985).
2. Vale, R.D. *et al. Cell* **43**, 623 (1985).
3. Hollenbeck, P. *Nature News and Views* **317**, 17 (1985).
4. Vale, R.D., Reese, T.S. & Sheetz, M.P. *Cell* **41**, 39 (1985).
5. Mitchison, T. & Kirschner, M. *Nature* **312**, 232 (1984).
6. Miller, R.H. & Lasek, R.J. *J. Cell Biol.* **101**, 2181 (1985).

Peter Hollenbeck is in the MRC Cell Biophysics Unit, 26-29 Drury Lane, London WC2B 5RL, UK.

Repeated sequences

The origin of retroposons

from John Rogers

THE chromosomal DNA of mammals is prolifically infested with dispersed repeated nucleotide sequences of various types, scattered through the non-coding regions of the genome. One of the most abundant and most mysterious types is the long interspersed nucleotide element or LINE class (see figure). Two recent adv-

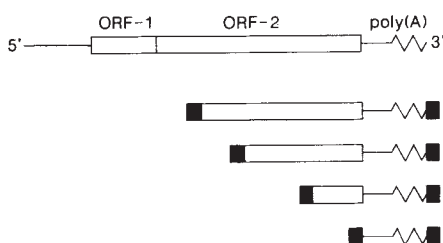


Diagram of the arrangement of a full-length LINE-1 sequence (top line) and of 5'-truncated dispersed repeats (lower lines) flanked by short direct repeats (black squares). ORF, open reading frame.

ances have produced an extraordinary new view of these repeated sequences.

Understanding of these sequences has been hindered by their complexity, but it is clear that they all belong to a single family (LINE-1), both in rodents and in primates, which has a consensus structure about 6 kilobases long, although most copies are randomly truncated 3' fragments¹⁻⁴. Most copies have a poly(A) (or A-rich) tail at the 3' end and are flanked by short repeats. This suggests that LINEs are inserted reverse transcripts (retroposons⁵), similar to processed pseudogenes and many other dispersed repeats. Their origin and operation remain to be elucidated, but the new developments allow these problems to be formulated with greater precision.

First, a full-length LINE-1 transcript has been discovered in a human teratocarcinoma cell line⁶. Heterogeneous LINE-1 transcripts are common in nuclear RNA,

but this is the first report of an abundant, full-length, 'plus'-strand, cytoplasmic LINE-1 RNA. Such a transcript is required by the reverse transcription theory of LINEs. Its presence in a teratocarcinoma cell line, which may represent the earliest cleavage stage of the embryo, and its down-regulation when these cells differentiate, are of special interest with respect to retroposon origins. A family of short retroposons in mice, called B2, is also abundantly transcribed in the early embryo⁷. And it is noteworthy that mammals, which begin transcribing the zygotic genome at the two-cell stage, contain numerous retroposon copies of many kinds of RNA, whereas other animals, which do not begin transcription until later, have few retroposons⁸. Therefore, the cleavage-stage embryo could be where new retroposons are made.

The second major advance is the compilation of full-length LINE-1 DNA sequences, which show that LINEs are derived from parental copies with long conserved open reading frames. Maxine Singer and colleagues have compiled a collection of random primate specimens into a 'pastiche' complete sequence, which has several open reading frames that could be joined by a few point mutations¹. Others have sequenced complete single specimens from human⁹ and mouse¹⁰; the human specimen has many interruptions in the reading frames but the mouse has two perfect open reading frames spanning 5 kilobases. This seems to be the arrangement from which other LINE-1 specimens have degenerated. Although most of them are defective in the open reading frames because of 5'-truncation, frameshifts and/or stop codons, comparisons of many specimens show clearly that the parental open reading frames have been evolving under selection for their coding potential since the mammalian radiation^{11,12}.

So what is the function of the open reading frames? And why have these genes

produced a thousand-fold more processed pseudogenes than have other genes? Two alternative theories can now be formulated. One possibility is that LINEs are 'selfish' genes in the manner of retroviruses or transposons, propagating via transcription and reverse transcription, and that the open reading frames encode proteins involved in this process. Indeed, in their size and arrangement, the two open reading frames (joined by a short overlap) resemble the *gag-pol* region of retroviruses⁹. Loeb *et al.*⁹ point out two other features of the full-length mouse LINE-1 sequence which enable them to construct a novel theory along these lines. First, there is a possible homology (on the borderline of significance) between the larger open reading frame and the reverse transcriptases of retroviruses. Second, there is an array of tandem repeats at the 5' end, 4 2/3 in one specimen and 1 2/3 in another, which Loeb *et al.* suggest may carry promoters. Stochastic expansions of this array could solve the major problem in propagating a gene through multiple rounds of transcription and reverse transcription, which is to regenerate the promoter at the 5' end.

However, these 5' repeats may be a recent acquisition in mouse LINE-1, as they are absent from primate LINE-1. In contrast, the coding sequences have been conserved in all mammalian orders, evolving in parallel with the whole organism¹². This is consistent with the alternative possibility, that there is a conserved parental LINE gene with a function for the organism. The vast number of retroposon copies could be explained if this gene normally undergoes reverse transcription as a physiological process, of which the dispersed copies are occasional by-products.

Both theories propose that LINE-1 RNA is the normal substrate for reverse transcription to form retroposons. If so, all the other retroposons in mammals may be by-products of the same machinery, which occasionally acts on the wrong substrate RNA. Fortunately, the complete DNA sequences not only permit these theories to be formulated, but also open up the possibility of testing them. □

1. Voliva, C.F. *et al. Nucleic Acids Res.* **11**, 8847 (1983).
2. Grimaldi, G., Skowronski, J. & Singer, M.F. *EMBO J.* **3**, 1753 (1984).
3. Rogers, J.H. *Nature News and Views* **306**, 113 (1983).
4. Singer, M.F. & Skowronski, J. *Trends biochem. Sci.* **10**, 119 (1985).
5. Rogers, J.H. *Int. Rev. Cytol.* **93**, 187 (1985).
6. Skowronski, J. & Singer, M.F. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6050 (1985).
7. Vasseur, M., Condamine, H. & Duprey, P. *EMBO J.* **4**, 1749 (1985).
8. Hattori, M., Hidaka, S. & Sakaki, Y. *Nucleic Acids Res.* **13**, 7813 (1985).
9. Loeb, D.D. *et al. Molec. cell. Biol.* **6**, 168 (1986).
10. Martin, S.L. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2308 (1984).
11. Hardies, S.C. *et al. Molec. Biol. Ecol.* (in the press).
12. Burton, F.H. *et al. J. molec. Biol.* **187**, 291 (1986).

John Rogers is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Quantum mechanics

Is the theory applicable to macroscopic objects?

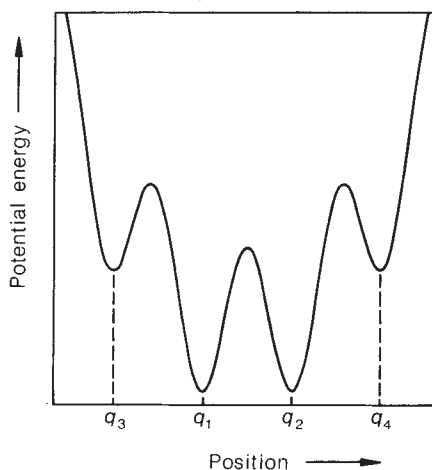
from Ulrich Eckern

DESPITE the fact that quantum mechanics has been used successfully for decades to explain physical processes on an atomic or sub-atomic level, a slight uneasiness has remained about the connection between quantum mechanics and daily life, which is governed by the laws of classical mechanics. This relation is unique: although quantum mechanics contains newtonian mechanics as a limiting case, at the same time it requires this limiting case for its own formulation. The importance of the concept of measurement, that is, the interaction of a quantum object with a classical object (a measuring apparatus) has been elucidated by Niels Bohr. Objections to these ideas have been formulated, perhaps most rigorously by Albert Einstein (see ref. 1). Naturally, the question arises of how quantum theory extrapolates to macroscopic systems: are there interference phenomena associated with the superposition of wavefunctions describing macroscopically distinct states? Is the transition from quantum to classical behaviour continuous or discontinuous? Clearly, these questions have to be answered experimentally.

To demonstrate the principle of superposition, let us consider an object which can be described by a single variable, say q (the coordinate of a 'particle'), with a potential energy $U(q)$ as shown in the figure. Evidently, the potential has two equivalent stable minima, q_1 and q_2 , and a classical particle in its ground state (the state of lowest energy) will be located at q_1 or q_2 . In the quantum case, because of the uncertainty principle, the energy of these two states is increased to the value $h\nu/2$ (zero point motion), where h is Planck's constant and ν the frequency of classical oscillations of small amplitude (near q_1 or q_2). In addition, the exponentially small overlap of the wavefunctions localized at q_1 and q_2 , respectively, leads to the possibility of coherent tunnelling between '1' and '2': The probability of finding the particle in state 1 at a certain time t , $P(t)$, is oscillating between one and zero (with a frequency much smaller than ν). This phenomenon, quantum coherence, is well established for microscopic systems — for example, the oscillations between the two positions of the nitrogen atom in an ammonia molecule (ammonia maser), and the oscillations between the two values of strangeness of a neutral K-meson.

What are the difficulties in applying quantum mechanics to macroscopic ob-

jects? Much theoretical²⁻⁷ and experimental⁸⁻¹² work has been inspired by the ideas of A.J. Leggett². The important point is that, even if the description of a macroscopic object by a single collective variable, say q , is a good approximation (and this can be justified for specific models⁴), there is always an environment, caused by all the microscopic degrees of freedom, leading to friction in the motion of this variable. Theoretical results⁵⁻⁷ indicate a behaviour that depends crucially on the value of the dimensionless friction constant, α , as well as on temperature, T . Especially, for $T = 0$, three regimes have to be distinguished. For $0 \leq \alpha < 1/2$, coherent oscillations



are obtained in $P(t)$ (quantum behaviour), whereas for $\alpha > 1$, the system is localized in one of the minima (classical behaviour). The intermediate range (except for $\alpha = 1/2$) is an unresolved problem⁷. In addition, for $\alpha < 1$ and at temperatures that are not too low (depending on α), and for all temperatures if $\alpha > 1$, incoherent relaxation is predicted, with $P(t)$ decreasing exponentially with time.

Relevant experiments⁸⁻¹² have been performed recently using superconducting quantum interference devices (SQUIDs) and Josephson junctions at very low temperatures (a few millidegrees from absolute zero). For example, the SQUID is a superconducting ring interrupted by a weak link, in which case the collective variable q is the magnetic flux through the ring, and oscillations of q correspond to an alternating current. The figure shows the potential relevant to this case, provided a certain external magnetic field is applied. In addition, α is equal to R_0/R , where R is the resistance of the SQUID and R_0 is close to

$h/4e^2 \approx 6,453$ ohms (e is the charge of an electron). But macroscopic quantum coherence has not yet been observed. Instead, experimental investigations have focused on the tunnel effect, a separate but related aspect of quantum mechanics.

Suppose that the particle is initially localized in an unstable minimum of the potential (at q_3 or q_4). The probability of finding the particle close to that minimum then decreases in time, and the decay time can be determined by monitoring the change of the magnetic flux. In agreement with the theoretical prediction³, it is found that the decay rate at low temperatures is larger than the rate predicted by thermal activation theory and approaches a temperature-independent value for $T \rightarrow 0$. This value is thought to be caused by quantum tunnelling. In addition, the $T = 0$ decay rate decreases with increasing dissipation, which agrees with the expectation that the width of a wavefunction (for example, a harmonic oscillator in its ground state) is reduced by contact with an environment.

However, experiments have gone one step further: if the decay rate out of the metastable state is small, then, for a quantum variable, the energy levels for states localized near the unstable minimum have to be quantized and, by irradiating the system with microwaves, the particle can be excited into higher states. Thus, the decay rate (which is larger for higher energy levels) has to increase if h times the microwave frequency equals the energy difference (of, for example, the transition from the lowest to the first excited level). This has been confirmed in an elegant experiment by J.M. Martinis, M.H. Devoret and J. Clarke¹⁰, which provides further evidence for the existence of macroscopic quantum phenomena. Although the idea of macroscopic quantum coherence still needs experimental justification, I believe that the experiments described here represent important steps in this direction, and that they may ultimately shed some light on the relationship between classical and quantum physics. □

1. Pais, A. *Subtle is the Lord... the Science and the Life of Albert Einstein* p.5 (Oxford University Press, New York, 1982).
2. Leggett, A.J. *Progr. Theor. Phys. (Suppl.)* **69**, 80 (1980).
3. Caldeira, A.O. & Leggett, A.J. *Phys. Rev. Lett.* **46**, 211 (1981).
4. Ambegaokar, V., Eckern, U. & Schön, G. *Phys. Rev. Lett.* **48**, 1745 (1982).
5. Chakravarty, S. *Phys. Rev. Lett.* **49**, 681 (1982).
6. Bray, A.J. & Moore, M.A. *Phys. Rev. Lett.* **49**, 1545 (1982).
7. Leggett, A.J. *et al. Phys. Rep.* (in the press).
8. Voss, R.F. & Webb, R.A. *Phys. Rev. Lett.* **47**, 265 (1981).
9. Washburn, S., Webb, R.A., Voss, R.F. & Faris, S.M. *Phys. Rev. Lett.* **54**, 2712 (1985).
10. Martinis, J.M., Devoret, M.H. & Clarke, J. *Phys. Rev. Lett.* **55**, 1543 (1985).
11. Schwartz, D.B., Sen, B., Archie, C.N. & Lukens, J.E. *Phys. Rev. Lett.* **55**, 1547 (1985).
12. Devoret, M.H., Martinis, J.M. & Clarke, J. *Phys. Rev. Lett.* **55**, 1908 (1985).

Ulrich Eckern is at the Institut für Theorie der Kondensierten Materie, Universität Karlsruhe, D-7500 Karlsruhe, West Germany.

Ice shelf creep rates and the flow law of ice

SIR—A suggestion¹ that creep measurements made on the Ward Hunt Ice Shelf, Canada, should provide needed information relevant to defining the flow law for ice, exposed the fact that such measurements had already been carried out². Attempts to improve our knowledge of the flow law for ice have been under way for the past 30 years and the latest attempt³ seems unlikely to be the last⁴.

This communication is intended to point out that the data referred to in ref. 2 cannot easily be interpreted to provide information relevant to investigations of the flow law for ice. I will attempt here to summarize the major difficulties surrounding this problem.

Whilst the arctic ice shelves and ice islands are potentially able to provide a low (stress, strain rate) point on deformation maps for ice, it is of crucial importance to carry out the experiments in the right place and at the right time. For two major reasons neither of the known experiments^{4,5} fulfill these conditions.

First, it has long been known that the internal structure of the arctic ice shelves and ice islands is extremely complex⁶, and may vary significantly over relatively short distances. Power reflection coefficient data from airborne radar measurements⁷ point to the existence of significant regional inhomogeneities in shelf structure. These variations occur in both crystallographic and chemical properties. In particular, two of the major ice types: sea ice, with large elongate crystals and iced firn with small equant crystals, may be contained in the same column, but in different proportions, depending on location. Crystal size, shape and optic axis orientation are all known to affect the creep properties of bulk ice⁸. Ice salinity may vary over a range of 15‰. The very significant effect on creep rate of the brine volume in the ice has recently been documented⁹. Thus, it is essential to site any experiment on flow law investigations at least where the structure of the shelf is known and preferably where it is least complicated.

Second, although most measurements of strain rate may be carried out with acceptable reliability, the determination of the accompanying stresses may not be, except in the simplest of cases¹⁰. The "point" at which the Ward Hunt Ice Shelf strain rate measurements¹ were made was < 1 km from the edge of an ice-locked island and < 4 km from the outer edge of the shelf. The geometry of the shelf in this region precludes a simple analytical stress solution. However, even if this problem were overcome (by finite element numerical modelling) yet another problem is present. The extremely mobile arctic pack ice impinges on the outer edge of Ward

Hunt Ice Shelf for the greater part of the year. There are periods when pack ice is thrust into and over shelf ice, as on ice island T3 (ref. 5). During such episodes, which may not be uncommon, significant compressive strains (and hence stresses) may be generated in shelf ice. During aftermath conditions on T3, an estimate of the compressive stress was $(0) - 0.01 \text{ MNm}^{-2}$ in the ice island, corresponding to $(0) - 0.1 \text{ MNm}^{-2}$ in the pack ice⁵. It is now evident⁹ that stresses in sea ice may exceed the latter value by an order of magnitude. Thus, for limited periods, the normal gravity expansive creep stresses (and accompanying strains) in thin ice shelves and ice islands, which are $(0) + 0.01 \text{ MNm}^{-2}$ (ref. 5) may be completely defeated, even though the long-term creep of the shelf ice may result in a net expansion⁴. Thus, the time span over which the flow law experiments are carried out is also of great importance. Then, one must evidently have to consider the relative effects of the different creep stages (primary, secondary and tertiary).

Experiments to simulate the creep of Ward Hunt Ice Shelf taking into account the structure and the sea ice back-pressure effect are now under way, using a finite element numerical model¹¹, and results shall be published shortly¹².

As some consolation to the array of difficulties presented here, it may be possible to avoid one of the major difficulties mentioned above. Part of the inner edge of Ward Hunt Ice Shelf flows in the form of a tongue into the entrance of an isolated fiord (Disraeli) which contains only passive fresh water ice. These conditions would seem to favour this as the best site for any flow law experiments. This leaves only the problem of the structural character of the shelf which, even if fully accounted for, does not allow easy comparison between these data the data³ for the much thicker Antarctic ice shelves of entirely different structure¹³.

G. HOLDSWORTH

*Environment Canada,
National Hydrology Research Institute,
101-4616 Valiant Drive NW, Calgary,
Alberta T3A 0X9, Canada*

Received 5 July; accepted 22 July 1985.

1. Weertman, J. *Nature* **314**, 227 (1985).
2. Hattersley-Smith, G. *Nature* **315**, 462 (1985).
3. Doake, C.S.M. & Wolff, E.W. *Nature* **314**, 255–57 (1985).
4. Dorrer, E.J. *Glaciol.* **10**, 211–25 (1971).
5. Holdsworth, G. & Traetteberg A. *Proc. Second Intern. Conf. on Port & Ocean Eng. Univ. of Iceland*, 419–440 (1973).
6. Jeffries, M.O. thesis, Univ. Calgary (1985).
7. Prager, B.T. thesis, Univ. British Columbia (1985).
8. Paterson, W.S.B. *The Physics of Glaciers 2nd Edn* (Pergamon, Oxford, 1981).
9. Sanderson, T.J.O. *Proc. 7th Int. Symp. Ice, Hamburg* **4**, 151–207 (1984).
10. Weertman, J. *J. Glaciol.* **3**, 38–42 (1957).
11. MacAyeal, D.R. & Thomas, R.H. *Annls Glaciol.* **3**, 189–194 (1982).
12. MacAyeal, D.R. & Holdsworth G. *J. geophys. Res.* (in the press).
13. Zumberge, J.H. & Swinbank, C. *Ant. Res. Ser. A.G.U. Monogr.* No. 7, 197–208 (1962).

AIDS virus and HTLV-I differ in codon choices

SIR—The genomic pattern of choices between synonymous codons ("coding strategy") represents a fundamental evolutionary structure useful for distinguishing species¹, and great evolutionary times have been shown to be associated with substantial changes in coding strategy^{2–4}. A rapid indication of the overall scheme of choices can be obtained by calculating the preferred triplet in each amino acid's set. Comparison between sequences or species is then possible by counting differences in preferred codons for the 18 amino acids with codon choices (methionine and tryptophan have unique codons except in mitochondria).

For instance, nuclear genes of mouse, rat and chicken have 16 to 18 preferences common with those of man. This is also an interesting index of the degree of imitation of host strategy by a virus or organelle. For example, when codon frequencies in the complete genome sequence of the AIDS (acquired immune deficiency syndrome) virus⁵ are compared to those of all sequenced human nuclear genes⁶, excepting the immune system, not a single common preferred triplet among the 18 cases is found. However, matching human T-lymphotropic virus type 1 (HTLV-I) preferences against those in man reveals 11 identical choices. Other retro-oncoviruses give similar results⁴. On the other hand, Visna lentivirus (VLV)⁷ shares no common codon preference with man. With both VLV and AIDS virus, the various codons in each amino acid set have greatly different relative frequencies from human genes⁶. Thus AIDS virus and VLV do not at all imitate human (or mammalian) codon usage while HTLV-I does, to a degree. (Not as much as some other viruses, however, as the complete genome sequences of EBV and Ad2 each show 15 common preferred triplets with human mRNA⁴.)

We now compare viruses among themselves. Perhaps surprisingly, HTLV-I and HTLV-III (AIDS virus) do not share any preferred codons, while VLV and AIDS virus coincide in all 18 cases. This suggests that the AIDS and Visna viruses have diverged relatively recently but that AIDS virus (or VLV) and HTLV-I are much more distantly related. Our results are compatible with those of Chiu *et al.*⁷ and Sonigo *et al.*⁸ who reported a close relationship in the genetic organization of VLV and AIDS virus. However, the present work concerns all three viral genes; in particular, we demonstrate that VLV-AIDS virus similarity is not limited to the *pol* gene and the common region between *pol* and *env*^{4,7,8}.

It does not necessarily follow that no individual genes in the host exhibit codon

choices similar to those in AIDS or Visna viruses. The pattern of preferred codons varies in and between different gene families¹⁻⁴; the widest variation observed so far is with immune system genes, in which even the constant and variable segments of the same immunoglobulin mRNA differ in this respect⁹. We believe that, in the context of a study of co-evolution of viruses and the immune system, an interesting comparison for AIDS/Visna virus coding strategy is that of different members of the threatened defense system. We find that some T-cell receptor coding sequences share triplet preferences with AIDS/Visna viruses, although the results remain difficult to interpret.

R. GRANTHAM
P. PERRIN

*Institut d'Evolution Moléculaire,
Université Lyon I,
69622 Villeurbanne, France*

1. Grantham, R. *et al. Nucleic Acids Res.* **8**, r49-r62 (1980).
2. Grantham, R. *Trends biochem. Sci.* **5**, 327-331 (1980).
3. Grantham, R. *et al. Nucleic Acids Res.* **9**, r43-r74 (1981); *Bull. Inst. Pasteur* **83**, 95-148 (1985).
4. Grantham, R. *et al. Oxford Surveys in Evolutionary Biology* **3** (in the press).
5. Ratner, L. *et al. Nature* **313**, 277-284 (1985).
6. GenBank release 35 (1985).
7. Sonigo, P. *et al. Cell* **42**, 369-382 (1985).
8. Chiu, I.M. *et al. Nature* **317**, 366-368 (1985).
9. Perrin, P. *Nucleic Acids Res.* **12**, 5515-5527 (1984).

Origin of the human AIDS virus

SIR—Recent reports¹⁻³ as well as recent correspondence to *Nature*⁴ have dealt with the possible animal origins of the human AIDS (acquired immune deficiency syndrome) virus. I would like to consider other theories in addition to cross-species transmission for the origin of human T-lymphotropic virus type III/lymphadenopathy-associated virus (HTLV-III/LAV).

One theory holds that HTLV-III/LAV has long been present in humans but confined to a small number of populations. Changing migration patterns (for example from rural to urban Africa) coupled with some practices of modern society (intravenous drug abuse and large numbers of homosexual partners) may have produced sufficient infection with HTLV-III/LAV to make AIDS a noticeable disease. The poor transmissibility of HTLV-III/LAV is consistent with this theory.

A second theory holds that HTLV-III/LAV is a new variant of a retrovirus that has always infected mankind. There are a number of human diseases for which a viral aetiology is as yet only suspected or for which the causative viral agent has not yet been identified. These include non-A non-B hepatitis, multiple sclerosis and rheumatoid arthritis. Some of the features of these diseases resemble those of lentivirus-induced disease. HTLV-III/LAV could conceivably be a new, T-cell tropic variant of a previously unidentified human lentivirus.

The recent isolation of a HTLV-III/LAV-related simian virus (called STLV-III) first from macaques⁵ and then from African green monkeys⁶ has raised speculation regarding a third theory—that HTLV-III/LAV is a human-adapted virus arising from cross-species transmission from monkeys. Such cross-species transmissions have been documented, although rarely, in virology. The most recent occurrence is the appearance of canine parvovirus as a new dog-adapted virus in 1978^{7,8}. However, the mere existence of HTLV-III/LAV-related viruses in monkeys provides no evidence for cross-species transmission to humans. In fact, the low degree of sequence conservation between HTLV-III/LAV and STLV-III (<75%) argues against the transmission of one of these monkey viruses to humans relatively recently in history. Such sequence divergence has certainly not occurred after cross-species transmission of feline pan-leukopenia virus to dogs⁹.

In the absence of a clear-cut candidate for cross-species transmission, we can expect arguments on the origin of the human AIDS virus to continue.

RONALD C. DESROSIERS

*Harvard Medical School,
New England Regional Primate
Research Center,
Southborough,
Massachusetts 01772, USA*

1. Kanki, P. *et al. Lancet* **i**, 1330-1332 (1985).
2. Wong-Staal, F. & Gallo, R. *Nature* **317**, 395-403 (1985).
3. Norman, C. *Science* **230**, 1140-1142 (1985).
4. Rosenoir, J. *Nature* **317**, 100 (1985).
5. Daniel, M.D. *et al. Science* **228**, 1201-1204 (1985).
6. Kanki, P. *et al. Science* **230**, 951-954 (1985).
7. Tratschin *et al. J. gen. Virol.* **61**, 33-41 (1985).
8. Parrish *et al. Science* **230**, 1046-1049 (1985).

The occurrence of reversible epithelia

SIR—A recent letter to *Nature*¹ described an interesting case of an epithelium, the morphological and functional polarity of which can be reversed experimentally. The accompanying *News and Views* article² described it as "a curious violation of the usual rules" that "offers a promising system for learning how epithelia become programmed during development".

In fact, this case is a representative example of what is likely to be a general property of epithelia. Morphological and functional reversal of epithelial polarity was first discovered by Mauchamp and his colleagues using thyroid follicles more than 10 years ago^{3,4} and studied further by Nitsch and Wollman⁵. Since then, a vast body of literature has accumulated about similar phenomena in other systems; to name but a few examples, the gut of the sea urchin embryo⁶, retinal pigmented epithelium⁷ and the epiblast of the early chick embryo⁸. The stimulus for reversal can be anything that produces an asymmetric environment across the tissue, includ-

ing pH, serum concentration, cell contact, salt concentration and electrical potential. The mechanism of reversal has received most detailed attention from Jaffe⁹.

It seems to me that the above systems, in which epithelial polarity can be manipulated experimentally using embryonic tissues, might be more suitable for studying the relevance of this phenomenon to embryonic development than the more recent case^{1,2}.

C. D. STERN

*Department of Human Anatomy,
South Parks Road,
Oxford OX1 3QX, UK*

1. Schwartz, G.J., Barasch, J. & Al-Awqati, Q. *Nature* **318**, (1985).
2. Diamond, J.M. *Nature* **318**, 311 (1985).
3. Mauchamp, J. & Fayet, G. *Endocr. exp.* **8**, 170 (1974).
4. Mauchamp, J. *et al. Cell Tissue Res.* **204**, 417 (1979).
5. Nitsch, L. & Wollman, S.H. *J. Cell Biol.* **86**, 875 (1980).
6. Amemiya, S. *et al. J. exp. Zool.* **210**, 177 (1979).
7. Crawford, B.J. *Tissue Cell* **15**, 993 (1983).
8. Stern, C.D. & MacKenzie, D.O. *J. Embryol. exp. Morph.* **77**, 73 (1983).
9. Jaffe, L.F. *Phil. Trans. R. Soc. B* **295**, 553 (1981).

Phenolic degradation by "dissimilatory plasmid"

SIR—We have been studying the degradation of phenolic substances by bacteria and we have examined the various terms used by investigators to designate the plasmids associated with their degradation. They have been termed "degradative plasmids"¹, "catabolic plasmids"² and "metabolic plasmids"^{3,4}.

The term "degradation" suggests the immediate product formation (transformation of the substrate) whereas dissimilation would mean the complete cleavage of a substrate leading to its utilization as an energy source. Most of the reports indicate the complete or near complete dissimilation of aromatic substances. Therefore in our opinion, the expression "dissimilatory plasmid" is more indicative of the functional role of the plasmid than the other terms suggested.

S. BALAJEE,
K. BOOMINATHAN
A. MAHADEVAN

*Centre for Advanced Study in Botany,
University of Madras,
Madras-600 025, India*

1. Chakrabarty, A.M. *Proc. natn. Acad. Sci.* **70**, 1641-1644 (1973).
2. Williams, P.A. *Trends biochem. Sci.* **6**, 23-26 (1981).
3. Wong, C.L. & Dunn, N.W. *Genet. Res. Camb.* **27**, 405-412 (1976).
4. Gunsalus, I.C. & Yen, K.M. in *Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids* (eds Levey, S.B. *et al.*) 499-509 (Plenum, New York, 1981).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a rather technical character which are not provoked by articles or letters previously published (where the Matters Arising section remains appropriate).

The end of a great experiment?

from Robert Walgate

A strong swing to the right is expected in next month's general elections in France. This could mean a decline in the political importance enjoyed by French scientists since 1981.

WILL the 16 March general elections in France mark the end of the great experiment launched by President Mitterrand to mobilize French science on behalf of the economy? Or a new realism and a new beginning? The question is open, though many doubt the commitment to a research and innovation policy by M. Jacques Chirac, the Mayor of Paris, who is the man most tipped to become France's next prime minister.

"It's been an exceptional period for science" says one senior researcher, "and extremely important for France . . . but it may not extend beyond 1988", when the next presidential elections take place (there is a seven-year term for the presidency but only five for the government). In 1988 too the current "three-year plan" for research, the second and most sophisticated of Mitterrand's acts on behalf of science and technology, comes to an end.

The decline in political interest could come much sooner. Chirac has promised to dismantle the present ministry of research and technology, dismember the major research council, the Centre National de la Recherche Scientifique (CNRS), and give back the specialized research councils for agriculture (INRA) and medicine (INSERM) to the ministries of agriculture and health. On the other hand, he has promised more power and autonomy to the universities, and the creation of a ministry for higher education and research. But his advisers seem unconcerned that research should contribute to innovation. Because there will be no cabinet minister responsible for the coordination of research, the political weight of science under the right-wing coalition expected after 16 March is likely to shrink.

That is not to say that everything achieved in France under the present regime is wonderful, or that research is about

to collapse under a right-wing government but it is certainly creditable that the Mitterrand government has taken French research and development spending from 1.8 per cent of gross national product (GNP) in 1981 to 2.3 per cent in 1985 (government figures). The government has also engineered an extraordinary about-turn in the attitudes of academics from hatred of "profit" and "industry" to a fondness for the former enemy. On the other hand, by the granting of jobs for life — civil service status — to French researchers, the government has not yet won the intended benefits of mobility. This decision may not have been a disaster, but it is now widely regarded as a fundamental mistake. (The critics may yet be confounded. Mobility shows signs of increasing. At present only about ½ per cent of government researchers move each year, whereas 7 per cent is thought healthy, but the past few months have seen transfers of CNRS personnel into industry for periods of a year or more steadily increasing.)

Doubts over the Mitterrand administrations' treatment of tenure are outweighed by other achievements. The chief of these is that the President has opened up a profound debate on French problems, and in many cases set the scene for the solution of some of them. One notable achievement is that *le Parti Socialiste*, only recently forged from many odd political elements, has been formed in its first five short years of power into such a radical source of ideas on the relationship between science and prosperity.

The challenge

At the outset, President Mitterrand and his science ministers, beginning with the dynamic Jean-Pierre Chevènement, had to deal with a demoralized scientific establishment which had never really recovered from the rapid expansion of university posts in the late 1960s and the indifference of the 1970s. The student revolution of 1968 may have persuaded the right-wing governments of the 1970s, beginning with that of President Pompidou, to regard the universities, and the sciences with them, as mere sources of dissent.

During the 1970s, science support declined, except in the big (and successful) programmes of space and nuclear power (which were led by engineers rather than

scientists). It began to pick up only in 1980 as President Giscard d'Estaing belatedly began to realize the importance of science in international economic competition.

Mitterrand's accession in May 1981 was therefore an opportunity for an expansionist science policy. He was able to deliver both money and, after a decade of



Jean-Pierre Chevènement: his Colloque Nationale set the experiment going.

indifference, a little political interest in the importance of science. And the Socialist Party which he had constructed during the 1970s was a pragmatic party. For the French socialists, "science criticism" à la Herbert Marcuse and the Frankfurt school was only marginal. Science for the Socialist Party was to be an objective tool of government economic policy, as well as an ideological necessity (science = objectivity = analysis = truth).

The socialists, in 1981, needed not only to reinvest in science but to open up the system. There is a tendency in France to "corporatism", the identification of individuals with their castellated and protected institutions: for example the university system and, separately, the *grande écoles*, or elite engineering schools that cream much of the academic talent from the schools but do little research; the research councils, from which only one third of university teachers receive grants; industry; the great goal-oriented programme agencies such as the space agency, the Centre National des Etudes Spatiales (CNES), and the atomic energy agency, the Commission à l'Energie Atomique (CEA). This divisiveness of the French system had to be attacked.

Moreover largely as a legacy of 1968 and the massive university recruitment in

Average annual rise in CNRS budgets, 1982–86

Application of results ("valorisation")	25%
Computer facilities	19%
Medium equipment (FF0.5–5.0 million)	14%
Heavy equipment (over FF30 million)*	6.1%
Special programmes ('ATP')	3%
Laboratory operating funds	1.8%
TOTAL	6.5%

Values exclude salaries and allow for net 29% inflation over the four years.

* Includes operating costs.

a left-wing atmosphere thereafter, followed by an effective block on promotions and appointments during the 1970s, universities became heavily trade-unionized and egalitarian. Another key problem, therefore, was to reintroduce a degree of selectivity and evaluation into the university system, and to make the universities less suspicious of outsiders — particularly industry.

The response

The "opening up" of this locked system was probably Mitterrand's greatest achievement in science policy. But the extraordinary and immediate effect of the 1981 elections was that suddenly a whole clutch of issues surrounding science in modern France began to be tackled with enormous vigour. The policies were driven both by the political importance that President Mitterrand had given the programme, and by the revolutionary zeal of M. Jean-Pierre Chevènement.

Chevènement began with a move since admired by his predecessor, Giscard d'Estaing's science supremo Pierre Aigrain: he called a series of regional colloquia, and a national gathering in Paris in 1982, to get the various "corporatist" forces talking to one another about the role of science and technology in France. Suddenly everyone found there was common ground, and undoubtedly since then the whole direction of science in France has changed. Before 1982, *pilotage par l'aval* — essentially the turning of research to applied ends, with a hint of industrial control — was deplored.

Laboratory operating funds per full-time CNRS researcher (thousands of FF 1985)

1976	129
1978	106
1980	97
1982	115
1984	107
1986	112

After a 25% decrease in the late 1970s, laboratory operating funds per researcher have grown 15% since 1980 but still not returned to their 1976 level. Data exclude computers, medium and heavy equipment.

Now it is embraced, though under cover of a slightly different concept — called the *réponse à la demande socio-économique*.

Before the Colloque Nationale, the image of a scientist in France was that of an intellectual unsullied by matters of profit and production. After the colloquium, perhaps because of it, science was suddenly not isolated, but at the service of society. The result is that science is enjoying a new "double confidence", according to a senior political adviser to the present French science minister, Hubert Curien: on the one hand the politicians have confidence in the scientists, and on the other the scientists have confidence in their new role in France.



Jacques Chirac, hoping to be the next prime minister, promises industrialists a "rupture" from socialism.

The new-found willingness to work with industry is in marked contrast with the situation just ten years ago, when CNRS researchers went on strike and demonstrated against the first contract then being drawn up between the research council and industry (in the shape of the chemical giant, Rhône Poulenc). Now CNRS has not just one, but 377 contracts with industry, worth some FF 56 million (£5 million) annually to the research council. And the research council has drawn up outline agreements for all kinds of future research agreements, with dozens of major companies.

Moreover, under new more liberal statutes granted to the research councils, it is now possible at last in over-regulated France for research bodies not only to set up joint research, but also to establish common laboratories with industry and even to create their own industrial companies to develop products derived from research. CNRS, for example, has so far created three companies — in robotics, metrology and publishing; INSERM, the medical research council, has established a company for developing diagnostic kits from monoclonal antibodies; and a number of councils together have created the genetic engineering company Transgène in Strasbourg.

The great *décloisonnement* — the opening up of research to the outside world, and the dismantling of the corporatist barriers — goes further: for the research councils now actually talk to one another and to the universities. INSERM, the medical council, and CNRS, for example, have rationalized the joint funding of many laboratories and now actually have a number of common committees. CNRS has also for the first time drawn up agreements with every university, so it can develop its laboratories but still respect the universities' fledgling autonomy.

Higher education repair

One of the greatest problems faced by the present government was the reform of the universities, and of their relation with the other great institutions of higher educa-

tion in France, the élite *grandes écoles*, which cream the top 1 per cent from the million students in French higher education. Politically, the *grandes écoles* dominate French government and industry, and could be said to be broadly on the right. Though small, the *grandes écoles* are highly regarded and influential; but they did little research and granted no research degrees. The universities, by contrast, are more on the political left, are generally despised (with a few exceptions) and appear to be powerless except *en masse*. They do not even have a strong influence on research policy — which is mostly controlled by the research councils — yet almost all government-funded research, even through the research councils, is done on university campuses.

It was not until 1984 that a new law, the *Loi Savary* (after the then minister of education), was drawn up and voted to replace the old (1968) constitution for French higher education. In itself, the law failed to change the situation greatly. Too many conflicting pressure groups were at work, and the law was a compromise, giving a few promotions to the solidified university lecturing cohort of the 1960s, but actually strengthening local egalitarianism in the universities and increasing the teaching load. The unbalanced relation between university, *grandes écoles* and research councils remained more or less untouched.

However, the law itself was not all. And as the legislators predicted at the time, the real changes came as the results of codicils to the law, and its later detailing. For example, the complicated system of four different higher degrees has now been swept away and replaced by a single, three-year Anglo-American style PhD which is granted not nationally like the older degrees but locally, bearing the stamp of the university that granted it. There is also in place now a structure for the evaluation of the performance of each university, chaired by the universities' strong but fair critic, mathematician Laurent Schwartz; his committee, established just last year, has now submitted its first

report (on Strasbourg); and universities are for the first time defining research priorities, as part of three-year plans agreed with the ministry of education. Moreover, the *grandes écoles* — to the alarm of some universities — are beginning to do more and more research, and this year 10 *grandes écoles* have been granted the right to award PhDs. There is also a promise — from a probably outgoing government, it must be admitted — to establish more universities of technology, like the highly successful University of Compiègne, which has remained unique in France precisely because it bridged the gap between the *grandes écoles* and the universities, and therefore threatened both.

In the more traditional universities, policy has concentrated on introducing greater selectivity within the system. However, the selectivity is still controlled centrally by the ministry of education in Paris, mainly through a series of “multi-annual contracts” for research which have now been signed between two-thirds of the 75 or so French universities and the ministry, with the rest on the way.

In these contracts the ministry guarantees research support and positions at a defined level to the universities for three years — regardless of political change — but only provided they concentrate resources in their areas of strength. These, in turn, are defined by independent, expert panels established by the ministry, but working with the university presidents and regional governments and local industry. The result, although the sums of money involved are so far small, is that some real, objective selection is underway within the university for perhaps the first time since 1968.

Here then are the threads of a really fundamental change in French higher education: a streamlined research system with evaluation within and without the universities, which will lead to greater power for university presidents over their internal egalitarian councils; and competing research in the *grandes écoles* which should also encourage these elite schools to make their student selection on more objective grounds than at present. Moreover the contracts that universities are signing with the research councils, with regional governments (whose research budgets have increased nearly three-fold since 1981) and local industry are also strengthening both their autonomy and responsibility.

Away from Paris

One of the great strengths of the present administration has been its emphasis on regional autonomy — surprising perhaps, for a socialist government might be expected to be centralist rather than devolutionary. In research policy, this has meant the establishment of highly active and in-

Average annual net growth in CNRS by field of research, 1982–86

Life sciences	6.1%
Engineering	5.6%
Chemistry	3.5%
Earth, space and astronomy	3.1%
Maths and physics (excluding nuclear)	2.6%
Humanities and social sciences	2.6%
Nuclear	-1.6%
TOTAL	3.1%

Figures include salaries and heavy equipment but exclude large computing facilities. Net 29% inflation assumed.

dependent regional programmes, centred on regional (“state”) governments. These, not only in obvious centres of research such as Grenoble, Toulouse and Strasbourg but also in other areas, have been able to develop real local research and development strengths, taking into account not only regional industrial interests but also local training policy and plans for regional development.

The regional policy has enjoyed considerable financial backing from the central government. For example, the ministry of education has helped establish the “*Poles FIRTECH*” which link the *grandes écoles*, universities and industry in a region around technology judged important for the future and relevant locally.

These “poles” — some 16 are being developed and 50 are projected by 1988 — are to focus on one of 50 technologies which the ministry of research has identified as crucial to industrial innovation, including very large scale integrated electronics (VLSI), computer-aided design, artificial intelligence, robotics and biotechnology. Regions, concentrating on their own existing scientific and industrial strengths (Toulouse is strong in biotechnology, space science and robotics, for example; Grenoble in electronics and solid state physics) then opt to act as the relevant *pole FIRTECH* for the technology. Locally, a management group coordinates all relevant efforts, such as the subjects of PhD theses (many supported by industry through a new joint government–industry grants scheme, the *bourses CIFRE*) and the courses taught at the local universities.

One central aim of the policy was to find a sufficiently attractive, challenging — and potentially profitable — focus for high fliers from the *grandes écoles*, who immediately on leaving their undergraduate courses can command far higher salaries in industry and in government administration than they could receive as research students. This is where the *bourses CIFRE* come in, for under this scheme a young *grande école* graduate is employed by a company, but government repays half his salary if he is sent on to do a PhD.

The grand programmes

Thus a few of the major elements of the present French government’s research

policy have developed. There are many others left unmentioned, like the national space programme in which France leads the rest of Europe — and deep sea research (Frenchmen have reached the bottom of the Japan trench in the vessel *Nautil*, and made investigations there, deeper than any other scientists have gone). The nuclear energy agency, the CEA, is largely a law unto itself and is pressing ahead with fast-breeder development at high speed, again leading Europe, and despite the reluctance of the national utility, Electricité de France.

These ambitious projects are proving expensive, and there are fears that the French aim to give Europe the capability to put man in space (the Hermes programme) could draw so much from the French research budget that by next year support for basic science may face cuts to compensate. (According to the scientific advisory committee to the research minister, a 2 per cent fall for basic science in 1987 is likely under current targets. The committee is now pushing for an annual financial



Michele Alliot: the ex-chef du cabinet at Alice Saunier-Seïte, universities minister under Giscard d'Estaing. She aims to give more power to the universities.

ceiling on the space programme.) One way of spreading the cost would be a strong European dimension for the space and other big programmes, hence the government’s enthusiasm for the Eureka programme of bilateral and multilateral agreements on high technology products.

CNRS reform

It is too soon to say whether Eureka will succeed or fail. Back home, though, there are other problems of scale. One is CNRS itself, the major French research council and the great guarantee of research quality in France since the 1940s. With 25,000 employees and a budget which now exceeds even that of the civil side of CEA, some say that CNRS has become too big. According to Michèle Alliot, Chirac’s spokeswoman for higher education and research, CNRS is slow, bureaucratic, and overburdened by the problem of staff management; and, she claims, its role in the universities is divisive.

Others feel at least that the body is too monolithic — it covers subjects from literature to high energy physics — and in need

of internal reorganization. There are three proposals in the air for the future of CNRS, at this politically uncertain time: to copy the Anglo-American solution, and have an agency like the US National Science Foundation principally distributing funds to the universities; to split CNRS into several institutes; and, less drastically, to make CNRS more flexible by creating research "networks" within the existing structure.

But all agree — including M. Chirac's advisers, who seek to strengthen the universities at the expense of bodies like CNRS and INSERM to create American-style university autonomy — that the universities are not yet ready for the first solution. "It would be a catastrophe . . . the end of research in France" to dismantle the research councils now, says one Strasbourg scientist, as the universities are too weak to shoulder the responsibility. As for the second solution, the splitting up of CNRS, the French tendency to "corporatism" could mean that the separate parts could become too independent and conflictual. Already in biology, for example, French scientists are subject to policy-making decisions at over half a dozen different institutions — CNRS, INSERM, INRA (agriculture), even the CEA, the Institut Pasteur, the Curie Institute, the faculties of medicine. There are already felt to be too many actors on a small stage.

The present government's preferred solution, and that of the CNRS directorate, is to "soften" CNRS by creating "networks": associations of groups and laboratories which would focus on particular topics for a limited period of time (perhaps five years) with a degree of autonomy and guaranteed support. Unlike dismemberment, this would not create new institutions or, it is hoped, the associated bureaucratic inertia, but would allow new structures to develop within the existing body.

Where next?

French research still faces many problems. For example, new posts have been created steadily in the research councils according to industrial and research needs, but in the universities appointments have responded solely to the needs of education. This means that some fashionable departments, such as informatics, have been growing too fast (by 10 per cent a year) whereas others such as basic physics are almost stagnant. The high population of technicians in France (a few years ago there were 2–3 per researcher, though now the figure is nearer 1) led the present government to neglect technician appointments, and neither have there been many promotions or training schemes. So now an ageing population of disaffected technicians, unfamiliar with new techniques, is somewhat belatedly being recognized as a major

problem facing French science.

The French right, now poised for power according to the opinion poles, would aim to dismantle the research councils and encourage much greater autonomy in the universities. Concurrently, the right would move to reduce egalitarianism and increase selectivity in research by challenging the power of the unions in university and research council committees.

A new law detailing such moves is being prepared jointly by the two main parties on the right (the RPR and the UDF), and if could be before parliament by the summer. Chirac, it is said, is promising no gradualism but a "rupture" from old ways. Chirac's advisers are also saying that the ministry of research and technology with its 500 staff is too big and bureaucratic, and that something like the old Délégation Générale à la Recherche Scientifique et Technique (DGRST) with 50 staff would be more flexible and effective. Run by an official with the ear of the prime minister, this body, it is claimed, could have as much power of coordination as the ministry. It would control the total government civil research budget. But those drawing up the new law do not appear to have much interest in innovation policy, at least as defined by the present government, and they would abandon plans to create new universities of technology.

Links between universities and industry would come about because universities would need to seek financial support from industry, says Gérard Milhaud, a Paris medical professor who Chirac has placed in charge of drawing up the new law. The civil service status for existing scientists would not be challenged, but new appointees will not be civil servants, says Alliot.

Research council scientists in the universities, who now have few teaching responsibilities, will be integrated into the university and given some teaching to do, "suppressing the antagonism between university and research councils which is very sad for this country", says Milhaud. But "It would be absolutely foolish to change the research councils if we did not radically change the universities". These should be based on a faculty structure, giving authority and responsibility to the persons in charge. This would not be a return to pre-1968 feudalism, the argument goes, as professors and laboratory directors would be selected by national peer review. "I have no objection to peer review and evaluation. One of the problems is that because of unionization we have no effective system of evaluation at the research councils or the universities", Milhaud says.

Milhaud argues that whereas many of the measures adopted by the present government are good, few of them go far enough. For example, the selective multi-annual grants to universities can support the best groups at only half the rate enjoyed by a CNRS laboratory. "If that's all

I had [Milhaud has both INSERM and CNRS grants] I'd have to go gambling." There has been a doubling if the number of inventions patented at CNRS since 1980 "but ask them how many of the patents are used . . .". It is not the business of researchers to seek patents and neither do they have the skills required. And in capitulating to the unions and in granting civil servant status, the present government has "spoiled everything".

However, Milhaud's solutions may not go down well in the universities, for the "social contract" of researchers with industry and the state was made by a predominantly left-wing university population with an at least nominally socialist government, and a largely nationalized industry. Will the "contract" survive the radical right-wing government that M. Chirac hopes to lead as prime minister? Many say not, despite the continued presence of President Mitterrand, who has two more years to run.

It is conceivable that some of the reforms proposed by the right, applied delicately to the already improving French research structure, could build on the success of the past five years' work in French science policy, and achieve results in certain areas that the socialists could not dream of. But the present course of French science is very new; and if the 16 March elections and post-election manoeuvres go M. Chirac's way (under Mitterrand's new system of proportional representation there must almost certainly be a coalition of parties), there will be more than a slight touch to the wheel.

M. Chirac promises a "rupture" from present policies, in science as in everything else. As for Hubert Curien, the unelected and apolitical technocrat who was appointed research minister two years ago, he has promised the less radical alternative of "evaluation". If M. Curien remains in post, CNRS and other organizations (even CEA) are to be studied and criticized in depth, with reports due later this year. In principle, Curien could stay even after a change of government — but, he said in Paris earlier this month: "I do have a certain solidarity with the present government. I could not remain if the changes in policy were too drastic."

Thus the choice for the French scientific electorate seems clear. The socialists would continue present policies, which are becoming increasingly selective and evaluative, but threaten a possibly overbearing increase in the scale of the "great programmes", particularly space; while the right promise "rupture" from all this intervention, and a careful handing back of power to the universities, but also imply a possible weakening of the political importance of science. □

Robert Walgate is the France correspondent of Nature.

VLBI limits on the proper motion of the 'core' of the superluminal quasar 3C345

N. Bartel*, T. A. Herring*, M. I. Ratner*, I. I. Shapiro* & B. E. Corey†

* Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA

† NERO, Haystack Observatory, Westford, Massachusetts 01886, USA

VLBI (very-long-baseline interferometry) observations between 1971 and 1983 have been used to determine the positions of the 'core' of the quasar 3C345 relative to the more distant compact quasar NRAO512 with a fractional uncertainty as small as 2 parts in 10^8 . The core of 3C345 appears stationary in right ascension to within $20 \text{ arc } \mu\text{s yr}^{-1}$, a subluminal bound corresponding to $0.7c$. The apparent velocities of the jets are superluminal, up to $14c$ in magnitude.

INVESTIGATION of the kinematics of celestial bodies has been central to astronomy throughout its history. Optical observations of stars have yielded determinations of proper motions with a precision of nearly $0.1 \text{ arc ms yr}^{-1}$ (ref. 1) and pulse-time-of-arrival² and radio link interferometer³ observations of radio pulsars have yielded proper motions with milliarcsecond per year precision. By contrast, measurements with very-long-baseline interferometry (VLBI), spanning a few years, can yield proper-motion determinations with uncertainties of only a few tens of microarcseconds per year, enabling proper-motion studies to be extended usefully to radio sources outside our Galaxy.

Since Hargrave and Ryle's⁴ detection of a compact radio component midway between the outer lobes of the strong radio galaxy Cygnus A (see ref. 5) and Whitney *et al.*'s⁶ discovery of apparent superluminal expansion of a quasar (see refs 7–9), astronomers have believed that active galactic nuclei and quasars consist of a stationary 'core' or centre of activity, and jets of material ejected from the core with relativistic velocity^{10–12}. The transverse component of such relativistic velocity may appear to exceed substantially the velocity of light when the ejection is directed close to the line-of-sight to the observer¹³.

An extensively studied example of an extragalactic radio source with an outer lobe of arcsecond size, and multiple components of milliarcsecond size, is the quasar 3C345. Several VLBI groups have measured increases in the angular separation between these components of up to $0.4 \text{ arc ms yr}^{-1}$ (see refs 14–18). Because 3C345 has a redshift of $z = 0.595$, corresponding to a proper-motion distance of 2.2 Gpc (Hubble's constant, $H_0 = 65 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = 0$), these increases in separation require apparent transverse velocity differences of up to $14c$. A key question arises^{19,20}: which components, if any, of the quasar are stationary and which are moving? Here we give the answer.

Precision astrometry

The celestial coordinates α and δ of any specific point in the brightness distribution of a radio source can be derived as follows from the time-dependent fringe phase ($0 \leq \phi(t) \leq 2\pi$) obtained from an interferometer operating at frequency ν :

$$\frac{\phi(t)}{2\pi\nu} = \tau(t) = \tau^g(t) + \tau^{\text{struc}}(t) + \tau^{\text{ion}}(t) + \tau^{\text{trop}}(t) + \tau^{\text{inst}}(t) + \tau^{\text{noise}}(t) + \frac{n(t)}{\nu} \quad (1)$$

where τ^g , τ^{struc} , τ^{ion} , τ^{trop} , τ^{inst} , τ^{noise} and n are, respectively, the contribution to the phase delay $\tau(t)$ from: (1) the geometry of the interferometer relative to a reference point in the radio source; (2) the source structure about that point; (3) the ionosphere; (4) the troposphere; (5) the instrumentation; (6) the thermal noise; and (7) the integer number of cycles of phase needed to resolve the phase ambiguity. For a discussion of the

precise meaning of t , see ref. 21. We next describe each term in turn and later show for our observations that all the terms can be determined with a precision sufficient for relative positions to be estimated with uncertainties as low as $50 \text{ arc } \mu\text{s}$ (see also ref. 20).

The term τ^g is given approximately by:

$$\tau^g(t) \sim \frac{B}{c} [\cos D \cos \delta \cos (A_0 + \Omega t - \alpha) + \sin D \sin \delta] \quad (2)$$

where B , A_0 , and D are the length, right ascension at epoch, and declination, respectively, of the baseline; c is the speed of light; and Ω is the angular velocity of the Earth. (The accurate expression²¹ includes aberration and relativistic effects.)

The term τ^{struc} can be written as

$$\tau^{\text{struc}}(t) \sim \frac{1}{2\pi\nu} \arg \left\{ \int B(x, y) e^{-i2\pi[u(t)x + v(t)y]} dx dy \right\} \quad (3)$$

where 'arg' indicates the argument of a complex number; $B(x, y)$ is the brightness distribution of the source; x, y are the angular distances in radians from the position α, δ of the reference point, in the directions of increasing right ascension and declination, respectively; and $u(t)$ and $v(t)$ are the components, in wavelength units, in the directions of increasing right ascension and declination, respectively, of the projection of the baseline vector onto a plane normal to the direction α, δ .

From dual-band observations yielding the phase delays $\tau_{\nu_1}(t)$ and $\tau_{\nu_2}(t)$ at radio frequencies ν_1 and ν_2 respectively, the ionospheric contribution $\tau_{\nu_1}^{\text{ion}}(t)$ to $\tau_{\nu_1}(t)$ at frequency ν_1 can be written as

$$\tau_{\nu_1}^{\text{ion}}(t) \sim \left[\left(\tau_{\nu_2}(t) - \tau_{\nu_2}^{\text{struc}}(t) \right) - \left(\tau_{\nu_1}(t) - \tau_{\nu_1}^{\text{struc}}(t) \right) \right] \left[\frac{\nu_1^2}{\nu_2^2} - 1 \right]^{-1} \quad (4)$$

for $\nu_1, \nu_2 \gg \nu_p$ where ν_p is the ionospheric plasma frequency. The uncertainty in our computation of τ^{ion} depends mainly on the error in our computation of the structure-phase contributions, $\tau_{\nu_1}^{\text{struc}}(t)$ and $\tau_{\nu_2}^{\text{struc}}(t)$, for the same reference point α, δ on the sky, that is, on the error in our 'registration' of the maps made at the two frequencies.

The term $\tau^{\text{trop}}(t)$ can be computed approximately from surface measurements of the pressure, temperature and humidity at each antenna site and a simple model of the troposphere²².

The term $\tau^{\text{inst}}(t)$ is defined as the phase delay in the instrumentation from the antenna feed to the point where the video signal is sampled for digital recording. It includes contributions from the behaviour of the frequency standards and the local oscillator chains and from polarization impurities²³ of the antenna-feed systems.

The term $\tau^{\text{noise}}(t)$ represents the corruption of $\phi(t)$ by noise, including amplifier noise and noise from the sky background.

Table 1 Observations

Session	VLBI system	Epoch of observation	Duration of experiment (h)	Frequency (GHz)	Antennas*
1	Mark I	1971.77	5.4	7.9	D, K
2	Mark I	1972.36	5.8	7.9	D, K
3	Mark I	1972.51	3.9	7.9	D, K
4	Mark I	1974.33	3.6	7.9	D, G, K
5	Mark III	1980.57	5.9	2.3/8.4	B, F, K, O, S, T
6	Mark III	1981.20	5.5	2.3/8.4	B, F, O, S
7	Mark III	1981.43	15.8	10.7	B, F, G†, K, O
8	Mark III	1981.45	10.5	2.3/8.4	E, F, K, O
9	Mark III	1982.96	12.0	2.3/8.4	E, F, K, O, S, T

The observations of session 7 were made in left-circular polarization, all other observations in right-circular polarization. A hydrogen-maser frequency standard was used at each antenna site to govern the local oscillator chain and the 'time-tagging' of the recorded data.

* B, 100-m-diameter antenna in Effelsberg near Bonn, FRG, belongs to the Max-Planck-Institut für Radioastronomie. From here onwards, we give for each antenna a representative value of the product, S_{sys} , of system temperature (K) and inverse sensitivity (Jy K^{-1}) for each of the frequencies in GHz: S_{sys} (2.3, 8.4, 10.7) = 85, 155, 120. D, 64-m-diameter antenna in Goldstone, California; belongs to NASA. S_{sys} (2.3, 7.9) = 29, 45. E, 18-m-diameter antenna in Westford Massachusetts; belongs to the Northeast Radio Observatory Corporation; S_{sys} (2.3, 8.4) = 3600, 6700. F, 25-m-diameter antenna in Ft Davis, Texas; belongs to Harvard College Observatory; S_{sys} (2.3, 8.4) = 2100, 3000. G, 43-m-diameter antenna in Green Bank, West Virginia; belongs to the National Radio Astronomy Observatory; S_{sys} (2.3, 7.9, 10.7) = 590, 500, 250. K, 37-m-diameter antenna in Westford, Massachusetts; belongs to the Northeast Radio Observatory Corporation; S_{sys} (2.3, 7.9, 8.4, 10.7) = 5800, 830, 830, 760. O, 40-m-diameter antenna near Big Pine, California; belongs to California Institute of Technology; S_{sys} (2.3, 8.4, 10.7) = 520, 1000, 290. S, 25-m-diameter antenna in Onsala, Sweden; belongs to Onsala Space Observatory; S_{sys} (2.3) = 1,000. T, 20-m-diameter antenna in Onsala, Sweden; belongs to Onsala Space Observatory, S_{sys} (8.4) = 4,000.

† Phase delays from two-element interferometers with G as one element were of exceptionally poor quality and, therefore, were not used for the astrometric analysis.

This term cannot be modelled, and thus it sets a limit on the astrometric accuracy achievable with any given set of observations.

The integer numbers $n(t)$ of cycles of phase can be determined iteratively through weighted least-squares analysis of the fringe-phase observables^{14,20} after the other terms have been estimated with sufficient accuracy by analysis of the (nearly) unambiguous group-delay observables, $d\tau/d(2\pi\nu)$, or the unambiguous phase-delay-rate observables, $d\tau/dt$, or both.

The astrometric error due to the contributions τ^g , τ^{ion} , τ^{trop} and τ^{inst} can be reduced considerably by observing a nearby source simultaneously or in an interleaved fashion. Thus, observations of the pair of sources 3C345 and NRAO512, separated on the sky by $\sim 0.5^\circ$, allow determination of their relative position with errors as small as 50 arc μs . (For precise determinations of relative positions from observations of other pairs of radio sources, see refs 24–27.)

Observations

We observed the pair of sources 3C345 (1641+399, $z=0.595$) and NRAO512 (1638+398, $z=1.67$; R. C. Lynds, personal communication) with the Mark I²⁸ and Mark III²⁹ VLBI systems in nine sessions spaced between 1971 and 1983 (see Table I). The Mark I observations in the 1970s and their reduction are detailed elsewhere²⁰. With the Mark III system, the pair of sources was observed alternately in a repeated cycle, usually 3C345 for 1–3 min and NRAO512 for 2–4 min, the latter source longer since its flux density was lower. After several cycles, an additional 4–7 min was allowed for tape change. The recorded data were correlated using the Mark III VLBI processor at the Haystack Observatory.

Maps

From the 3C345 data obtained in the 1970s, brightness distribution models, composed of two gaussian components, have been published previously¹⁴. In addition, from the data obtained in 1974, a crude map was produced³⁰. Because of the sparseness

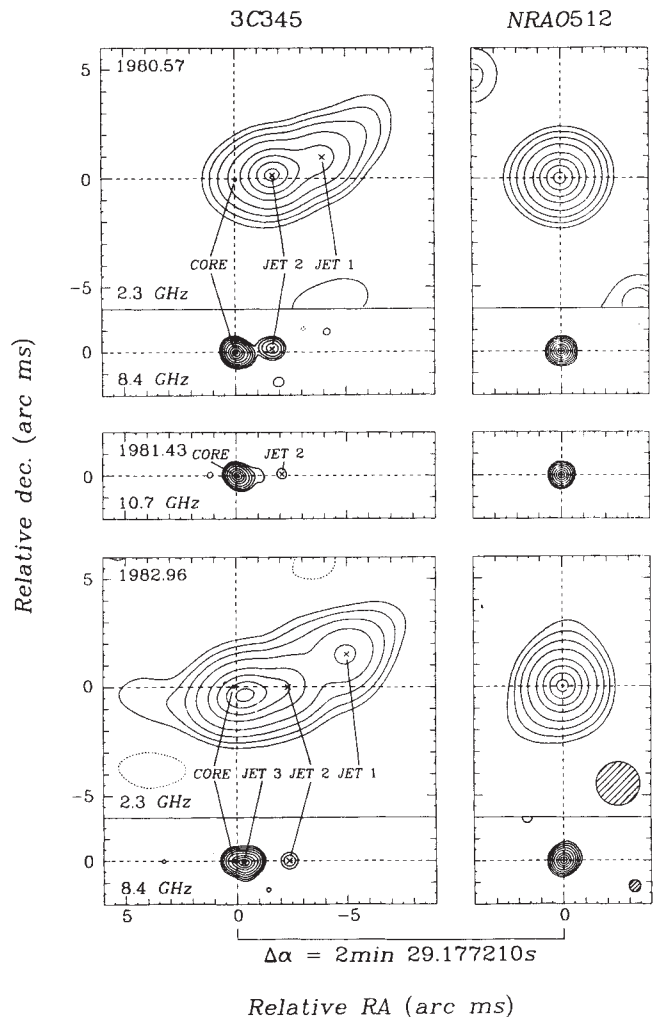


Fig. 1 Hybrid maps of 3C345 and NRAO512 at three epochs. The contours are at 95, 80, 60, 40, 20, 10, 5, 2.5, and -2.5% (dashed) of the peak brightness for each map and also at 1.25% and -1.25% in the 8.4-GHz maps. The restoring beam size (FWHM of a circular gaussian) is 2.0 arc ms for the 2.3-GHz maps and 0.55 arc ms for the 8.4- and 10.7-GHz maps (see the lower right-hand corner of the NRAO512 maps for epoch 1982.96) and is approximately equal to the width of the inner portion of the 'dirty' beam in the core-jet direction. All maps are shown to the same scale. The origins of all maps of 3C345 lie at the weighted mean of the positions of the brightness peaks of its easternmost (core) component relative to the position of the brightest peak of NRAO512, determined from the high-frequency data of the nine observing sessions. The positions of the core (●) and jets (×) were obtained from maps 'overresolved' by a factor of ~ 2 (see text).

of the u - v coverage of these early observations, we did not attempt to improve the quality of these models or map. For our observations since 1980, we calibrated the visibility amplitudes³¹ and used them together with the closure phases to produce with the Caltech software package³² a hybrid map of each source for each radio frequency band observed. For session 6 the limited u - v coverage did not allow us to make maps with a quality comparable to that of the other maps, and, therefore, we did not use the maps from this session. Instead, as the visibility data were similar to those from session 5, we used the brightness distributions from session 5 to calculate corresponding structure-phase corrections for session 6 (see Fig. 1).

To identify clearly a fixed reference point for 3C345, we 'overresolved' the maps about two-fold in the core-jet direction by convolving the set of δ -function (CLEAN) components of

each of the 8.4- and 10.7-GHz maps with a narrow circular gaussian beam of full-width at half-maximum (FWHM) 0.25 arc ms; we defined the resultant position of the brightness peak of the easternmost component, believed to be the 'core', as our reference point. In each of these maps, the second component, labelled Jet 2, is also visible; its location changes progressively from ~ 1.7 arc ms west of the core to ~ 2.5 arc ms west between 1980.57 and 1982.96. In addition, the map for epoch 1981.43 shows that the structure around the core is elongated towards the NE-SW direction, indicative of a new strong component (Jet 3) very close to the core. In our map for epoch 1982.96, this jet^{17,18} appears at a position located ~ 0.4 arc ms west of the core.

For the 2.3-GHz maps, a coarser resolution, and expected optical depth effects in the core region, led us not to follow the procedure used to choose a reference point for the maps made at higher frequencies. We again overresolved the maps two fold, but this time by convolving the set of δ -function components with a beam of width 1.0 arc ms (FWHM); we also marked the clearly identifiable peaks of Jet 1 and Jet 2, and selected a reference point for which the distance and direction to the peak of Jet 2 were equal to the corresponding quantities on the maps made at the higher frequencies. As jets usually have steep spectra in the gigahertz frequency range, the emission region of Jet 2 is expected to be optically thin between 2.3 and 8.4 GHz. Therefore, we assume that the brightness peaks of Jet 2 at the two frequencies coincide within an uncertainty of ~ 100 to 200 arc μ s, which depends on the quality of our maps. Our reference points at 2.3 and 8.4 GHz should coincide within the same uncertainty. In all maps of NRAO512, we chose as reference points the peaks of the brightness distributions of the (only slightly resolved) single component.

We later checked our choice for the registration of the brightness distributions at the 2.3- and 8.4-GHz frequencies by computing the apparent ionospheric delay as a function of time using equation (4). We had previously aligned the brightest components in the maps of 3C345 at these two frequencies and found an unexpectedly large and consistent disagreement of ~ 14 ps at 8.4 GHz for observations with the BO interferometer between the apparent ionospheric delay for 3C345 and that for NRAO512. Most of the disagreement was removed by adopting the registration described above.

Relative source positions

To estimate the proper motions of the components of 3C345 seen in the above maps, we analysed the phase delays, $\tau(t)$ (see equation (1)), to determine the position of the reference point in each map of 3C345 relative to the position of the reference point in the corresponding map of NRAO512. For sessions 1 to 4, we used the integral phase-ambiguity determinations $n(t)$ and the structure-phase corrections of Shapiro *et al.*²⁰ but changed the reference point from their choice of the centre of the (then brighter) jet to the core. For sessions 6 and 7, for each source and frequency band, we estimated $n(t)$ with our VLBI-3 program³³, corrected the unambiguous phase delays for structure phase and, for session 6 only, largely freed them from ionospheric contributions, as described in equation (4). For these six experiments we formed a differenced phase delay by subtracting from each phase-delay value for a 3C345 observation the corresponding value for the next NRAO512 observation. The order of the sources in the pairing of the observations was chosen to minimize the time differences between the paired observations.

For each observing session, we analysed with the VLBI-3 program the differenced phase delays, and the undifferenced delays for 3C345 only (as in ref. 20), to produce weighted-least-squares estimates of the reference position (J2000.0) in 3C345 relative to the reference position in NRAO512, along with coefficients of a fifth-order polynomial used to represent the clock behaviour at each site relative to that at the 'reference'

Table 2 Fixed parameters

Coordinates (J2000.0) for NRAO512*				
$\alpha = 16 \text{ h } 40 \text{ min } 29.632330 \text{ s}$ $\delta = 39^\circ 46' 46.029296''$				
Antenna site coordinates and electrical path lengths of troposphere				
Antennas†	Right-handed Cartesian coordinates‡			Tropospheric delay at zenith§ (ns)
	X (m)	Y (m)	Z (m)	
B	4,033,949.356	486,989.422	4,900,430.863	7.5
D	-2,353,619.320	-4,641,342.440	3,677,052.850	7.1
E	1,492,208.556	-4,458,131.334	4,296,015.883	7.8
F	-1,324,209.107	-5,332,024.067	3,232,118.941	6.6
G	882,881.833	-4,924,493.165	3,944,131.089	7.2
K	1,492,406.690	-4,457,267.330	4,296,882.102	7.8
O	-2,409,598.845	-4,478,350.477	3,838,603.741	6.9
S	3,370,968.054	711,464.922	5,349,664.140	7.8
T	3,370,607.962	711,916.453	5,349,830.868	7.8
UT1 and pole position				
See text		Earth tides¶		
		Radial Love number, $h = 0.610 \pm 0.003$		
		Horizontal Love number, $l = 0.085 \pm 0.008$		
		Tidal lag angle, $\theta = 0.0 \pm 0.3$ deg		

* The standard deviations for α and δ are < 1 arc ms (see text). The coordinate system in which the position of NRAO512 is expressed is identical to that described in ref. 42 except that their right ascension origin is offset eastward by 0.18 s from ours.

† See Table 1 for definition of abbreviations.

‡ The standard deviations of all components of all separations among the antenna pairs were < 4 cm.

§ The standard deviations for the delays at each of the antenna sites for sessions 1 to 4 and at sites B and S for sessions 6 and 7 were conservatively chosen to be 0.3 ns. In all other cases, the recording of surface meteorological data reduced the standard deviation to 0.1 ns.

|| The standard deviations for sessions 1 to 4 were ~ 2 ms for UT1 and ~ 20 arc ms for each component of the Earth's pole position (R. W. King, personal communication); for sessions 6 and 7, ~ 0.5 ms (for UT1) and ~ 5 arc ms (for the pole position); and for sessions 5, 8 and 9, ~ 0.1 ms (for UT1) and ~ 1 arc ms (for the pole position), respectively. The last set of values is considerably smaller because the relevant quantities were obtained from geodetic VLBI observations preceding our observations by only 1 day.

¶ The values are from a numerical computation by Wahr⁴³. The standard deviations are obtained by Herring⁴⁴ from the analysis of geodetic VLBI observations; his values are consistent with Wahr's.

site. The differenced phase delays were weighted equally, as were the undifferenced ones, but the standard deviation for the latter was increased 10-fold to match approximately the relative sensitivity of these two sets of observables to the coefficients of the clock polynomials and to the unmodelled parts of the tropospheric and, for session 7, of the ionospheric contributions to the delays, while still allowing the undifferenced data to determine the instrumental and atmospheric contributions.

For the remaining sessions (5, 8 and 9), scheduled in association with geodetic Mark III VLBI observations, we analysed the data using the Mark III geodetic analysis software³⁴, which is equivalent to the VLBI-3 program. The analyses themselves proceeded as for session 6, except for minor differences that do not significantly affect our final results.

The coordinates of the reference source NRAO512, of the antenna sites, and of the Earth's orientation, as well as the values for tropospheric zenith delays and for the Earth-tide parameters (see Table 2) were kept fixed in the weighted least-squares analysis. The first two sets of coordinates were determined by one of us (T.A.H.) from a large set of dual-band Mark III geodetic VLBI observations made between September 1980 and June 1983, except for the coordinates of antenna D (see Table 1 for meaning of abbreviation), which were tied to those of O through analysis of additional geodetic observations (J. W. Ryan and A. E. Niell, personal communications). The time-dependent, Earth-orientation parameters (UT1-UTC and the two components of the position of the Earth's rotation pole), all based on a consistent origin, were determined for sessions 5, 8 and 9 directly from geodetic VLBI observations within a few days of our astrometric observations. At other epochs, we relied on a

combination (see ref. 35) of results from several techniques: VLBI, satellite laser ranging, lunar laser ranging, Doppler tracking of navigation satellites and classical astrometry. The tropospheric zenith delay was computed as indicated above.

Our results for the coordinates of the easternmost component of 3C345, relative to the single compact component of NRAO512, are given for each of the nine sessions in Table 3 and are plotted in Fig. 2.

Error analysis

What is the uncertainty of our astrometric results caused by errors in the identification of the core in each of our maps of 3C345? For sessions 1 to 4, Wittels *et al.*'s¹⁴ determinations of separations suggest an uncertainty of ~ 100 arc μ s in both α and δ in locating the core component. For the other sessions, based on the complexity of the brightness distribution of the sources, the synthesized beamwidth and the dynamic range of the maps, we estimate that the uncertainty in α and δ stemming from our attempted identification of the core is ~ 35 arc μ s each for sessions 5, 7 and 9, and ~ 70 arc μ s each for sessions 6 and 8. Of course, an inadequate map of the source structure for any session also results in an error with an unknown diurnal time dependence in the structure-phase corrections to the phase delays. However, based on our trial use of plausible alternative maps, we find that the effect of such time-dependent errors on our final results, for sessions 5–9, is much smaller than the core identification uncertainties discussed above.

How much could the ionosphere have affected our astrometric results for the single-band observations of sessions 1 to 4 at 7.9 GHz and of session 7 at 10.7 GHz? From the sessions with dual-band observations at 2.3 and 8.4 GHz, we found that at 8.4 GHz, ionospheric effects on the results were as large as ~ 80 arc μ s and had, for α and δ separately, root-mean-square (r.m.s.) values of ~ 50 arc μ s. Thus, we assume the contribution of the ionosphere to the standard errors in our results is 60 arc μ s for sessions 1 to 4 and 30 arc μ s for session 7, even though for sessions 1 to 4 the ionosphere was probably not as active as it was for sessions 5 to 9.

How large are the errors of the results caused by uncertainties of the values of the parameters held fixed? We varied these values, one by one, by ~ 1 s.d. (see Table 2) and noted the difference introduced in our position estimates. The uncertainties in relative right ascension ($\Delta\alpha \equiv \alpha_{3C345} - \alpha_{NRAO512}$), and in relative declination ($\Delta\delta \equiv \delta_{3C345} - \delta_{NRAO512}$), resulting from the standard deviations of the values used for the fixed parameters, are listed for session 9 in Table 3.

The uncertainties in $\Delta\alpha$ and $\Delta\delta$ introduced by the standard deviations of the values used for the fixed parameters were added in quadrature with the corresponding standard deviations due to: (1) the ionospheric effects (only for sessions with single-band observations); (2) the core identification uncertainties; and (3) the statistical uncertainties, after the latter were (uniformly) scaled for each session so that the χ^2 per degree of freedom of the postfit residuals of the differenced phase delays was unity. The total standard deviation is given in Table 3 for each session. The smallest standard deviations are only two parts in 10^8 of the 0.5° separation of the two sources.

Proper motions

We determined the average proper motion, (μ_α, μ_δ) , of the core component of 3C345 by making, separately for $\Delta\alpha$ and $\Delta\delta$, a weighted least-squares fit to the data in Table 3. We find:

$$\mu_\alpha = -12 \pm 8 \text{ arc } \mu\text{s yr}^{-1}$$

$$\mu_\delta = -28 \pm 21 \text{ arc } \mu\text{s yr}^{-1}$$

The standard errors are the smallest ever obtained for the proper motion of any source. Corresponding $\sim 85\%$ (asymmetrical) confidence limit bounds on μ are $|\mu_\alpha| \leq 20$ arc $\mu\text{s yr}^{-1}$ and $|\mu_\delta| \leq 50$ arc $\mu\text{s yr}^{-1}$. Assuming $H_0 = 65 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0$,

Table 3 Estimates of the position of the core of 3C345 with respect to the core of NRAO512

Session	Epoch of observation	$\Delta\alpha - 2 \text{ min } 29.177210 \text{ s}^*$ (μ s)	(arc μ s)	$\Delta\delta - 1^\circ 50.96581''^*$ (arc μ s)
1	1971.77	16 ± 12	190 ± 140	580 ± 600
2	1972.36	6 ± 18	70 ± 200	410 ± 460
3	1972.51	38 ± 13	440 ± 150	$1,040 \pm 590$
4	1974.33	0 ± 13	0 ± 150	80 ± 420
5	1980.57	-4.1 ± 3.9	-47 ± 45	-27 ± 69
6	1981.20	-4.5 ± 7.1	-52 ± 82	-130 ± 170
7	1981.43	-6.1 ± 4.4	-70 ± 50	-30 ± 110
8	1981.45	-8.2 ± 6.8	-94 ± 78	100 ± 100
9	1982.96	6.8 ± 3.4	78 ± 39	-1 ± 49

For almost all sessions, the largest contribution to the standard deviation for $\Delta\alpha$ is the standard deviation of the core identification. This contribution ranges from 1.2 to 2.3 times larger than any other single contribution for each of the sessions except the second, for which it is only 0.9 times as large as the scaled statistical standard deviation. For $\Delta\delta$, the standard deviation contributions from standard deviations of the core identification, ionospheric delay, UT1, and atmospheric zenith delay, and from the statistical standard deviations, are more evenly divided. Negligible contributions come from the standard deviations of the pole position, of the values for the Earth tides, of the coordinates of NRAO512, and of the antenna site coordinates, even if, for the last, we consider the effect of continental drift, which is believed to be moving the North American sites with respect to the European ones at a rate of $\sim 2 \text{ cm yr}^{-1}$. The smallest statistical contributions to the standard deviations in $\Delta\alpha$ and $\Delta\delta$ were for session 7: 0.7 μ s for $\Delta\alpha$ and 17 arc μ s for $\Delta\delta$. The smallest total standard deviation was obtained for session 9; the standard deviations for $\Delta\alpha$ and $\Delta\delta$, respectively, are composed of the contributions from the core identification standard deviation (3.0 μ s, 35 arc μ s), from the statistical standard deviation (1.3 μ s, 22 arc μ s), and from the standard deviations of the atmospheric delay (0.7 μ s, 19 arc μ s), of UT1 (0.3 μ s, 18 arc μ s), and of the coordinates of the pole position, antenna sites and NRAO512 ($\leq 0.1 \mu$ s, ≤ 1 arc μ s).

* $\Delta\alpha \equiv \alpha_{3C345} - \alpha_{NRAO512}$, $\Delta\delta \equiv \delta_{3C345} - \delta_{NRAO512}$. The subtracted constants are the weighted averages, over all sessions, for the α and δ coordinates of the relative positions. All coordinates are in the J2000 system.

we infer relative velocity components in units of c of

$$v_\alpha = -0.4 \pm 0.3$$

$$v_\delta = -1.0 \pm 0.7$$

where v_α is nearly along the line connecting the core and jets.

We believe these proper motions are not significantly different from zero. Note that if we delete session 3, we obtain $\mu_\alpha = 0 \pm 10$ arc $\mu\text{s yr}^{-1}$ and $\mu_\delta = -17 \pm 23$ arc $\mu\text{s yr}^{-1}$, equivalent to $v_\alpha = 0.0 \pm 0.3$ and $v_\delta = -0.6 \pm 0.8$. These results constrain models^{36,37} (see also refs 38, 39) allowing or predicting significant proper motion of each visible component of a superluminal source.

In contrast to the core, the jets are clearly moving (see Fig. 3). The proper motion of the jet visible in the early to mid-1970s is ~ -160 arc $\mu\text{s yr}^{-1}$ in α , equivalent to $v_\alpha \sim -6$. Between 1975 and 1977, this jet appears to have accelerated to ~ -300 arc $\mu\text{s yr}^{-1}$, equivalent to $v_\alpha \sim -10$. Later, from 1977.6 onwards, the source appeared to possess two jets. The nature of the change is not clear. Perhaps the jet bifurcated or a new jet appeared; or perhaps the change was due only to the change in the technique used to determine source brightness distributions: modelling before 1977.5 and hybrid mapping thereafter. For two epochs we can compare the results from the use of these two techniques: (1) the hybrid map³⁰ made from the data from 1974.3 (session 4) shows one core and two jets, with the inner jet being at approximately the same position as the single jet in the two-component model¹⁴; and (2) a hybrid map⁴⁰ from epoch 1977.6 (close to the time of bifurcation) exhibits the same agreement with the model⁴¹ made from observations just two months earlier. Thus, the outer jet may have existed before 1976.5, but with insufficient flux density to be reliably modelled. The apparent motion of Jet 2 towards the core between 1977.5 and 1979.0 would in this context still be puzzling. In any event, we plotted from epoch 1977.6 onwards only jet positions obtained from hybrid maps. Since 1981 a third jet has been visible, and the proper motion of the outermost one has been ~ 400 arc $\mu\text{s yr}^{-1}$, equivalent to $v_\alpha \sim 14$.

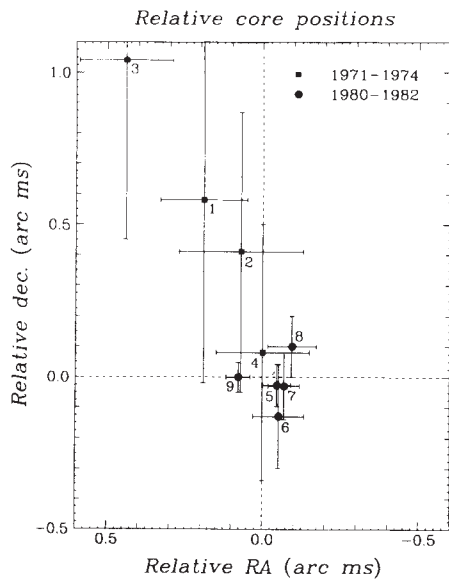


Fig. 2 The positions of the brightness peaks of the easternmost (core) component of 3C345 relative to the position of the brightness peak of NRAO512. The origin of the diagram lies at the weighted mean of the relative-position estimates. The numbers refer to the sessions listed in Table 1. For a description of the standard errors, see text.

Jitter of the core's position ?

The apparent stationarity of the easternmost component of 3C345 is consistent with this component being associated with the quasar's centre of activity and centre of mass. However, any detection of a significant change of the position of this component would not necessarily imply motion of the centre of mass²⁰. For example, the weighted mean position of this component from sessions 5 to 8 is 0.16 ± 0.03 arc ms west and 0.02 ± 0.05 arc ms south of the weighted mean position from sessions 1 to 4 and 9 and may indicate a slight shift of the brightness peak caused by the formation of the new component that was not clearly distinguishable from the core in sessions 5 to 8, but became distinguishable in session 9 as Jet 3, being located ~ 0.4 arc ms west of the core. A shift of the position of the peak of the brightness of the core may, therefore, indicate a very early stage of activity in the quasar's core region, which might be detectable with our astrometric techniques even before it becomes apparent in the hybrid maps.

Distance to the quasars

Our astrometrical results can be used to place 'plausibility' constraints (see also ref. 20) on 'local' theories of quasars. From the bounds on the relative annual parallax of 3C345 and NRAO512, we can place lower bounds on the distances to these quasars of, respectively, 15 and 20 kpc (85% confidence limits), if the quasars are assumed to be at sufficiently different distances from the Earth²⁰. These bounds are the largest yet obtained for parallax distance. Considering the direction to these quasars ($b^{\text{II}} \sim 41^\circ$, $l^{\text{II}} \sim 63^\circ$), note that these bounds would place the quasars outside the Galaxy. Furthermore, if we assume that the radial velocity of each quasar is nonrelativistic, that the transverse velocity v_{trans} of each is at least 100 km s^{-1} , and that the directions of these two velocities are randomly distributed, then the upper bound ($\sim 85\%$ confidence) on proper motion of $\sim 50 \text{ arc } \mu\text{s yr}^{-1}$ in any direction would place each quasar at a distance $> 200 \text{ kpc}$ ($> 70\%$ confidence) from the Earth. If, however, the redshift of each quasar is caused only by its velocity, and if v_{trans} of each is at least 20% of its velocity ($v_{\text{trans}} \geq 0.09c$ for 3C345 and $\geq 0.15c$ for NRAO512), then, by the same logic, the 'local' distance would have to be $> 40 \text{ Mpc}$ for 3C345 and 50 Mpc for NRAO512.

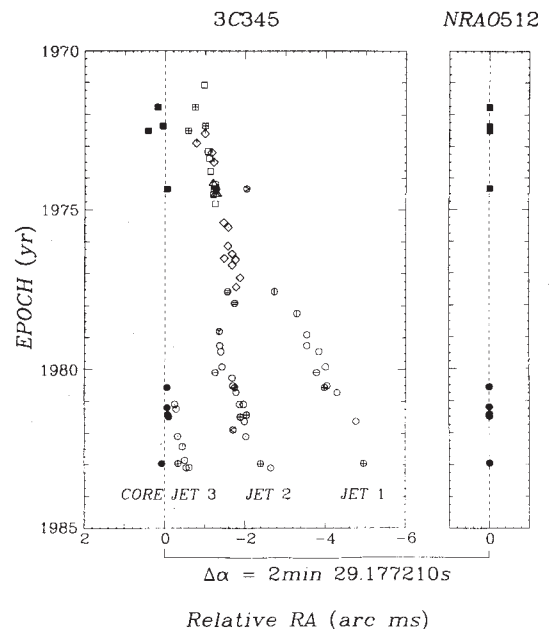


Fig. 3 The differences in right ascension between the core component (■, ●) and jet components (□, ⊕) in 3C345 and the single compact component in NRAO512 for each of the nine epochs. The dashed line marks the mean of these differences for the core component. Note that the core and jets are aligned nearly east-west. Additional symbols mark positions of the jet components relative to the dashed line. Round symbols are used for positions determined from hybrid maps. The data, with the observing frequencies in GHz in parentheses, are from: □, (7.9)¹⁴; ⊕, (10.7)⁴⁵; △, (10.7)⁴⁶; ◇, (10.7)⁴⁷; ◇, (5.0, 10.7)⁴¹; ⊕, (7.9)³⁰; ⊕, (5.0, 10.7)⁴⁰; ⊕, (10.7)⁴⁸; ⊕, (22.2)⁴⁹; ○, (5.0, 10.7)¹⁶; ⊕, (5.0)⁵⁰; ⊕, (2.3)⁵¹; ○, (10.7)¹⁷; ○, (22)¹⁸.

Proper motions of galaxies

VLBI astrometry, applied here to the study of quasars, can also be used for the determination of proper motions of galaxies that have compact radio cores. The nearby galaxy M81, at a distance of $\sim 3 \text{ Mpc}$, is a promising target because its radio brightness distribution is very compact. A decade of observations should yield a proper motion estimate with an uncertainty equivalent to a transverse velocity of $\sim 100 \text{ km s}^{-1}$.

We thank the Goddard-Haystack-MIT Mark III development team, and especially T. A. Clark, H. F. Hinteregger, C. A. Knight, A. C. E. Rogers, N. R. Vandenberg and A. R. Whitney for contributions. We also thank R. W. King for provision of data on UT1 and pole position, E. Bernstein, J. C. Breidenthal, G. J. F. J. Cuvelier and J. L. Davis for help with mapping of quasars, C. R. Gwinn for discussions and the staffs of the participating observatories for help with the measurements. N.B. thanks the Deutsche Forschungsgemeinschaft for partial support in an early stage of this project. N.B., B.E.C., M.I.R. and I.I.S. were supported in part by the NSF under grants nos AST-8300796 and PHY-8409671. T.A.H. and I.I.S. were supported in part by the Air Force Geophysics Laboratory Contract F19628-83-K-0031 and by NASA Contract NAS5-27571. The Haystack Observatory is operated by the Northeast Radio Observatory Corporation with support from the NSF. Astronomical VLBI at Harvard's G.R. Agassiz Radio Astronomy Station and at the Owens Valley Radio Observatory is supported by the NSF. The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the NSF.

Received 16 October; accepted 20 December 1985.

1. Russell, J. L., Gatewood, G. D. & Worek, T. F. *Astr. J.* **87**, 428-432 (1982).
2. Davis, M. M., Taylor, J. H., Weisberg, J. M. & Backer, D. C. *Nature* **315**, 547-550 (1985).
3. Lyne, A. G., Anderson, B. & Salter, M. J. *Mon. Not. R. astr. Soc.* **201**, 503-520 (1982).
4. Hargrave, P. J. & Ryle, M. *Mon. Not. R. astr. Soc.* **166**, 305-327 (1974).

5. Perley, R. A., Dreher, J. W. & Cowan, J. J. *Astrophys. J.* **285**, L35-L38 (1984).
6. Whitney, A. R. *et al. Science* **173**, 225-230 (1971).
7. Knight, C. A. *et al. Science* **172**, 52-54 (1971).
8. Cohen, M. H. *et al. Astrophys. J.* **170**, 207-217 (1971).
9. Porcas, R. W. in *Active Galactic Nuclei* (ed. Dyson, J. E.) 20-49 (Manchester University Press, 1985).
10. Scheuer, P. A. G. *Mon. Not. R. astr. Soc.* **166**, 513-528 (1974).
11. Blandford, R. D. & Königl, A. *Astrophys. J.* **232**, 34-48 (1979).
12. Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. Mod. Phys.* **56**, 255-351 (1984).
13. Rees, M. J. *Nature* **211**, 468-470 (1966).
14. Wittels, J. J. *et al. Astron. J.* **81**, 933-945 (1976).
15. Schraml, J. *et al. Astrophys. J.* **251**, L57-L59 (1981).
16. Unwin, S. C. *et al. Astrophys. J.* **271**, 536-550 (1983).
17. Biretta, J. A., Cohen, M. H., Unwin, S. C. & Pauliny-Toth, I. I. K. *Nature* **306**, 42-44 (1985).
18. Moore, R. L., Readhead, A. C. S. & Bååth, L. *Nature* **306**, 44-46 (1985).
19. Bartel, N. *et al.* in 'VLBI and Compact Radio Sources', *IAU Symp.* No. 110, 113-116 (1984).
20. Shapiro, I. I. *et al. Astr. J.* **84**, 1459-1469 (1979).
21. Shapiro, I. I. in *Methods of Experimental Physics*, Vol. 12, Part C, (ed. Meeks, M. L.) 261-276 (Academic, New York, 1976).
22. Davis, J. L., Herring, T. A., Shapiro, I. I., Rogers, A. E. E. & Elgered, G. *Radio Sci.* **20**, 1593-1607 (1985).
23. Bartel, N. *et al. Astr. J.* **90**, 2532-2539 (1985).
24. Marcaide, J. M. & Shapiro, I. I. *Astr. J.* **88**, 1133-1137 (1983).
25. Gorenstein, M. V. *et al. Astrophys. J.* **287**, 538-548 (1984).
26. Bartel, N. *et al. Astr. J.* **90**, 318-325 (1985).
27. Morabito, D. D. *Astr. J.* **90**, 1004-1006 (1985).
28. Whitney, A. R. *et al. Radio Sci.* **11**, 421-432 (1976).
29. Rogers, A. E. E. *et al. Science* **219**, 51-54 (1983).
30. Cotton, W. D. in *Radio Interferometry Techniques for Geodesy*, 193-197 (NASA Conf. Publ. No. 2115, 1980).
31. Bartel, N. *et al. Astrophys. J.* **262**, 556-563 (1982).
32. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **223**, 25-36 (1978).
33. Robertson, D. S. thesis, Massachusetts Institute of Technology, Cambridge, MA (1975). Reprint Goddard Report No. X-922-77-228 (Goddard Space Flight Center, Greenbelt, Maryland).
34. Clark, T. A. *et al. IEEE Trans. GARS GE-23*, 438-449 (1985).
35. Morgan, P. J., King, R. W. & Shapiro, I. I. *J. geophys. Res.* **90**, 12645-12652 (1985).
36. Chitre, S. M. & Narlikar, J. V. *Mon. Not. R. astr. Soc.* **187**, 655-659 (1979).
37. Bahcall, J. N. & Milgrom, M. *Astrophys. J.* **236**, 24-42 (1980).
38. Blandford, R. D., McKee, C. F. & Rees, M. J. *Nature* **267**, 211-216 (1977).
39. Marscher, A. P. & Scott, J. S. *Publ. astr. Soc. Pacif.* **92**, 127-133 (1980).
40. Readhead, A. C. S., Pearson, T. J., Cohen, M. H., Ewing, M. S. & Moffet, A. T. *Astrophys. J.* **231**, 299-306 (1979).
41. Seielstad, G. A. *et al. Astrophys. J.* **229**, 53-72 (1979).
42. Ma, C. *et al.* Preprint, Goddard Space Flight Center, Greenbelt, Maryland.
43. Wahr, J. M. *Geophys. J. R. astr. Soc.* **64**, 677-704 (1981).
44. Herring, T. A. in *Space Geodesy and Geodynamics* (eds Anderson, A. J. & Cazenave, A.) (Academic, New York, in the press).
45. Cohen, M. H. *et al. Nature* **268**, 405-409 (1977).
46. Shaffer, D. B. *et al. Astrophys. J.* **218**, 353-360 (1977).
47. Cohen, M. H. *et al. Nature* **268**, 405-409 (1977).
48. Cohen, M. H. *et al. Astrophys. J.* **247**, 774-779 (1981).
49. Bååth, L. B. *et al. Astrophys. J.* **243**, L123-L126 (1981).
50. Spencer, J. H., Johnston, K. J., Pauliny-Toth, I. I. K. & Witzel, A. *Astrophys. J.* **251**, L61-L63 (1981).
51. Cohen, M. H. *et al. Astrophys. J.* **272**, 383-389 (1983).

Activation of meiosis and sporulation by repression of the *RME1* product in yeast

Aaron P. Mitchell & Ira Herskowitz

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143, USA

The ability of the yeast Saccharomyces cerevisiae to enter meiosis and sporulate is conferred by the a1-α2 regulatory activity of the mating-type locus. We have cloned the RME1 gene and show that: the RME1 transcript is repressed by a1-α2 in a/α cells; overexpression of the RME1 product in a/α cells blocks sporulation; and inactivation of RME1 by a null mutation constructed in vitro permits cells lacking a1-α2 activity to sporulate. The RME1 product thus inhibits meiosis; the mating-type locus permits entry into meiosis by repression of RME1.

THE mating-type locus (*MAT*) regulates two transitions in the life cycle of the yeast *Saccharomyces cerevisiae*: the mating of haploids to form a diploid; and the sporulation of a diploid to yield haploid meiotic products. Cells that differ at *MAT* but which are otherwise genetically identical display fundamentally different properties (see refs 1, 2 for reviews). Cells with either *a* or *α* information at *MAT* are specialized for mating: *a* and *α* haploids respond to the mating factor secreted by cells of the opposite type (*α* and *a*, respectively) and undergo cell and nuclear fusion to produce an *a/α* diploid. The *a/α* cell, in contrast, is uniquely capable of meiosis and sporulation. These processes are initiated only upon nutritional limitation, which then sets in motion pre-meiotic DNA replication, high-level recombination, the two meiotic divisions and packaging of the four meiotic products into ascospores (see ref. 3 for review). The mating-type locus thus specifies states permissive for either mating or sporulation; these processes are initiated by the appropriate type of cell in response to environmental triggers (presence of mating factors or starvation).

Here we seek to understand how the *MAT* locus determines the ability of cells to sporulate. Sporulation requires that cells be *a/α* diploids: *a/a* and *α/α* cells are unable to sporulate⁴. Kassir and Simchen found that a mutation in the *RME1* (regulator of meiosis) gene bypassed this *MAT* requirement⁵: diploids homozygous for the *rme1-1* mutation sporulate regardless of their mating type. Subsequent studies led to the proposal that the *RME1* product is an inhibitor of sporulation and that *a/α* cells sporulate by virtue of turning *RME1* activity off in some manner; *a/a* and *α/α* diploids fail to sporulate because they express *RME1* activity⁶.

How might *MAT* regulate *RME1* activity? There is a group of genes whose transcripts are expressed in haploids and in *a/a*

and *α/α* diploids but repressed in *a/α* diploids. These haploid-specific genes include *MATa1* (refs 7, 8), *HO* (ref. 9), *KAR1* (M. Rose and G. R. Fink, personal communication), *STE5* (V. MacKay, J. Thorner and K. A. Nasmyth, personal communication) and the Ty transposon family¹⁰. Repression requires both the *a1* protein, encoded by the *MATa* allele, and the *α2* protein, encoded by *MATα*, which make up *a1-α2* activity⁷⁻⁹. *a1-α2* activity is also required for sporulation^{5,11}. It has been proposed⁶ that *a/α* cells are specialized for sporulation because *a1-α2* activity represses transcription of *RME1*, in other words, that *RME1* is a member of the set of haploid-specific genes. To evaluate this hypothesis, we have cloned the *RME1* gene and used this DNA to examine both natural *RME1* expression and the consequences of altered expression of the gene.

Isolation of *RME1*

We cloned the *RME1* gene by screening for a plasmid that could reverse the phenotype of an *rme1* mutation in an appropriate recipient strain. The recipient was a diploid whose ability to sporulate was dependent on being defective in the *RME1* gene (genotype *mata1/MATα rme1/rme1*). Transformants carrying plasmids from a yeast genomic library were tested for their ability to sporulate by a simple plate assay¹²: the recipient diploid was heterozygous for mutations conferring resistance to two drugs, canavanine and cycloheximide (genotype *can1/+ cyh2/+*), but was sensitive to the drugs because the resistance mutations are recessive. After meiosis, the diploid yields haploid segregants which carry both resistance mutations (*can1 cyh2*) and therefore grow in the presence of both drugs. Transformants that had acquired a copy of the wild-type *RME1* gene became unable to sporulate and were detected by their failure to yield drug-resistant progeny.

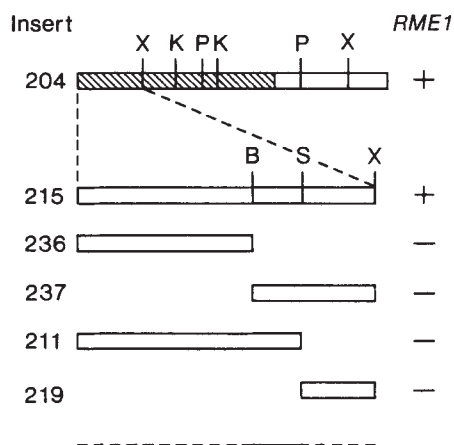


Fig. 1 Plasmids carrying *RME1* and their derivatives. The ability of each insert, carried in the low copy vector YCp50 (C. Mann, personal communication), to complement an *rme1* mutation is indicated in the right-hand column: +, complementation; -, no complementation. The inferred extent of the *RME1* gene is indicated by the solid line at the bottom of the figure, with uncertainty indicated by the broken line. The shaded area of insert 204 represents sequences shared by the four independent *RME1*-bearing plasmids described in the text. The sites recognized by restriction enzymes *XhoI* (X), *KpnI* (K), *PvuII* (P), *BglII* (B) and *SphI* (S) are indicated. Inserts are drawn to scale; insert 215 is 3 kb. Plasmids were introduced into strain AM-R103 (*mata1/MATa ura3/ura3 rme1/rme1 can1/+ cyh2/+* and heterozygous for other recessive auxotrophies) by spheroplast transformation¹³. Transformants, selected for *ura3* complementation, were purified (or replated for single colonies, when libraries were screened) on Synthetic Complete plates lacking uracil (SC-uracil plates)³⁶, then transferred to sporulation plates³⁶ for 5 days at room temperature. Sporulation ability was determined by replica-plating to SC-uracil-arginine+canavanine+cycloheximide plates, growth indicating the completion of meiosis by the plasmid-bearing cells in the colony, and by visualizing spores in asci upon microscopic examination of the cells suspended from the sporulation plates. The failure of a transformant to sporulate was determined to be plasmid-dependent by confirming that loss of the plasmid restored sporulation ability, as follows: transformants were patched onto non-selective YEPD plates³⁶ and transferred through three sequential outgrowths by replica-plating. This permitted the accumulation of cells that had lost the plasmid. The patches were then transferred to sporulation plates, incubated for 5 days at room temperature and finally replica-plated to SC-uracil-arginine+canavanine+cycloheximide and to SC-arginine+canavanine+cycloheximide. Absence of growth on the first plate confirms that cells containing the plasmid do not sporulate; presence of growth on the second plate demonstrates that loss of the plasmid allows cells to sporulate. Insert 204 is one of the original *RME1*-bearing isolates from which all the other inserts shown were derived. A *ClaI* site in the vector, just to the left of the shaded region of insert 204, was used in transferring inserts 215, 236 and 211 into YCp50; these were ligated into the vector *ClaI* site at one end and into the *SalI*, *BamHI* and *SphI* sites, respectively, at their other ends. Inserts 237 and 219 were carried between the *BamHI* and *SalI* or *SphI* and *SalI* vector sites, respectively.

The *RME1* gene was isolated from a library (M. Rose, J. Thomas, P. Novick and D. Botstein, personal communication) of yeast DNA inserts generated by partial digestion with the restriction endonuclease *Sau3A* and carried in the low copy-number *URA3* vector YCp50 (C. Mann, personal communication). Transformants, selected for complementation of a *ura3* mutation, were transferred to sporulation plates by replica-plating and incubated for 5 days to allow sporulation. The colonies were then assayed for production of drug-resistant progeny by replica-plating onto medium containing canavanine and cycloheximide. Colonies that failed to yield drug-resistant growth were picked from the sporulation plates and examined microscopically for the absence of ascospores. We then

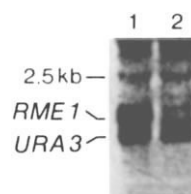


Fig. 2 Identification of the *RME1* transcript. RNA was isolated⁹ from exponential YEPD cultures of isogenic strains 246-1-1 (*MATa RME1 leu2 ura3 trp1 his4 can1*, lane 1) and AM1220-1B (*MATa rme1-2 leu2 ura3 trp1 his4 can1*, lane 2), fractionated by electrophoresis on an agarose-formaldehyde gel³⁷, and transferred to nitrocellulose paper³⁷. The paper was hybridized¹⁷ with the nick-translated probes pAM225, which contains insert 215 (Fig. 1) carried in plasmid pBR322 (ref. 38), and YIp5 (ref. 39), which carries the yeast *URA3* gene. The paper was then washed and exposed to Kodak XAR-2 film at -70 °C with a Dupont 1 Lightning intensifying screen. Strains carrying *rme1* mutations, such as AM1220-1B, were distinguished from *RME1* strains by their ability to form sporulating diploids when mated to testers of genotype *HMLa mata1-1 HMRa sir3-8 rme1*.

examined 153 such colonies to determine whether their inability to sporulate was caused by the presence of the plasmid: loss of the plasmid allowed sporulation for 20 transformants (see Fig. 1 legend).

Four independently isolated plasmids were compared by restriction enzyme digestion. The restriction map of one of the smaller inserts, from plasmid pAM204, is shown in the upper part of Fig. 1. This insert shared 9 kilobase pairs (kb) of homology with the three other plasmids (indicated by the shaded portion of the map). We anticipated that this segment would include the *RME1* gene.

We confirmed that this segment contains the *RME1* gene by demonstrating that it can direct plasmid integration¹³ to the chromosomal *RME1* locus. A restriction fragment of the pAM204 insert, shown in Fig. 1 as insert 215, was used to construct the integrating plasmid pAM227L, which also contains the yeast *LEU2* gene. This plasmid was transformed into strain AM1205-9B (*MATa leu2 RME1 ade6*), and integration was promoted¹⁴ by cutting the plasmid at the unique *SphI* site in the insert. Two *Leu*⁺ transformants were crossed to AM1200-1C (*MATa leu2 rme1 ADE6*) and meiotic tetrads were analysed to map the plasmid integration site. We found that 32 of 33 tetrads displayed 2 *Leu*⁺*Rme*⁺ : 2 *Leu*⁻*Rme*⁻ segregation, indicating that the plasmid had integrated at *RME1*. This was verified by the linkage of the integration site to its centromere and to *ADE6* (ref. 6). (The exceptional tetrad contained four *Leu*⁻*Rme*⁻*Ade*⁺ spores and therefore presumably arose from a diploid cell that had become homozygous for one arm of chromosome 7 before sporulation.) We conclude that these inserts, which provide *RME1* function, do indeed contain the *RME1* gene. J. Margolske and G. Simchen (personal communication) have also recently cloned the *RME1* gene.

Identification of the *RME1* transcript

The *RME1* transcript was identified in three steps. First, the approximate boundaries of the gene were ascertained by complementation tests. Next, an insertion mutation in the *RME1* gene was constructed *in vitro* and introduced into the genome. Finally, transcripts homologous with the *RME1* region were compared in the insertion mutant and wild-type strains.

The *RME1* gene boundaries were estimated by testing fragments for the ability to provide *RME1* function. Insert 215 (Fig. 1), a 3-kb *Sau3A/XhoI* fragment, complemented the *rme1* defect. Deletions of insert 215 from either side to the internal *BglII* and *SphI* sites abolished complementation (inserts 211, 219, 236 and 237; Fig. 1). *RME1* activity thus requires that these *BglII* and *SphI* sites are intact but does not depend on information to the right of the *XhoI* site at one end of the 215 insert.

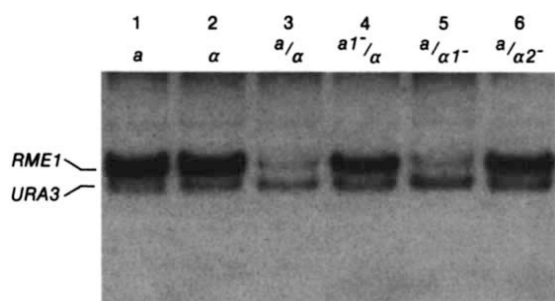


Fig. 3 Control of *RME1* transcript levels by the *MAT* locus. *RME1* transcript levels were measured in haploids and diploids bearing the indicated *MAT* alleles by hybridization to filter-bound RNAs that had been fractionated by electrophoresis, as described in Fig. 2 legend. The hybridization probes were the internal 0.6-kb *Bgl*III/*Sph*I fragment of insert 215 (Fig. 1) carried in the vector pUC18 (ref. 40) along with the *URA3* probe Ylp5. The strains used were isogenic to strain EG123 except as indicated, and the *mat* mutations³⁵ were provided by P. Siliciano. Haploid strains: EG123 (*MATa leu2 ura3 trp1 his4 can1*, lane 1), 246-1-1 (*MATα*, lane 2); diploid strains: AM1788 (*MATa/MATα*, lane 3), AM1789 (*mata1-50/MATα*, lane 4), AM1790 (*MATa/mata1-189*, lane 5) and AM1791 (*MATa/mata2-182*, lane 6).

We constructed *rme1* insertion mutations by ligating a 3-kb *Bgl*III fragment carrying the *LEU2* gene into the *Bgl*III site of insert 215 in two orientations. These mutant alleles were then used to replace wild-type genomic sequences¹⁵ by transformation into the *MATa/MATα leu2/leu2 RME1/RME1* diploid AM1788. Southern blot analysis^{16,17} confirmed that several of the *Leu*⁺ transformants were heterozygous for a 3-kb insertion at the *Bgl*III site in the *RME1* region (data not shown). Three such diploids were sporulated and yielded 2 *Leu*⁺:2 *Leu*⁻ progeny in 23 of 23 tetrads, demonstrating that insertion had occurred at a single locus. All the *Leu*⁺ progeny failed to complement an *rme1* tester, thus confirming that these insertions had created *rme1* mutations. We designate these insertions the *rme1-2* and *rme1-3* alleles, in which the *LEU2* gene is transcribed away from or towards the *Sph*I site in *RME1*, respectively.

An *rme1-2* strain served as the source of RNA lacking the normal *RME1* transcript. Northern blots of RNA from wild-type strains contain transcripts of ~2.5 and 1.3 kb that hybridize to *RME1* DNA (Fig. 2, lane 1). The *rme1-2* mutation abolishes production of the 1.3-kb species (lane 2), as we have also found for the *rme1-3* mutation (data not shown). Furthermore, this 1.3-kb transcript traverses the *Bgl*III and *Sph*I sites that were inferred to be internal to *RME1* but does not extend beyond the *Xho*I site that defined one endpoint of the gene (see Fig. 1; data not shown). These results establish that the 1.3-kb RNA is the *RME1* transcript.

Regulation of *RME1*

It has been proposed⁶ that the *MAT* gene products *a1* and *α2* confer sporulation ability by negative regulation of *RME1* activity. Previous work has demonstrated that the set of haploid-specific genes is transcriptionally repressed by *a1-α2* (refs 7-9). Therefore, we used the cloned *RME1* gene as a probe to test whether regulation of *RME1* by *MAT* is exerted at the level of transcript accumulation (Fig. 3). Both *a* and *α* haploid cells produce similar levels of the 1.3-kb *RME1* transcript (Fig. 3; lanes 1, 2). The isogenic *a/α* diploid, in contrast, produces much less of this transcript (Fig. 3; lane 3). Densitometer tracings of this blot, using the *URA3* transcript as a normalization standard, indicate that the *RME1* transcript level is reduced 10-20-fold in *a/α* diploids compared with its level in haploid cells. This is a much smaller repression ratio than has been found for the other haploid-specific genes *HO* (ref. 9) and *MATα1* (refs 7, 8). We confirmed that the *HO* transcript is repressed to a much

Table 1 An *rme1* null mutation bypasses the *a1-α2* requirement for sporulation

<i>MAT</i>	% Sporulation	
	<i>RME1/rme1-2</i>	<i>rme1-2/rme1-2</i>
<i>a/α</i>	68	75
<i>a1⁻/α</i>	<0.2	45
<i>a/α2⁻</i>	<0.2	25

Diploids with the *MAT* and *RME1* alleles shown above were grown on YEPD plates for 2 days at 30 °C, then transferred to sporulation plates and incubated at 23 °C for 10-14 days. The frequency of asci (% sporulation) in each population was determined by microscopic examination of at least 100 cells. Each frequency is the average obtained with two independent diploids issuing from each mating. The parents of the *a/α* diploids were AM1228-15C (*MATa rme1-2 leu1 ura3 trp1 can1*), AM1220-1A (*MATα RME1 leu2 ura3 trp1 his4 can1*) and AM1220-1B (*MATα rme1-2 leu2 ura3 trp1 his4 can1*). The parents of the *a1⁻/α* diploids were AM1228-3C (*MATα rme1-2 can1*), AM1229-4D (*mata1-50 RME1 leu2 ura3 trp1 his4 can1*) and AM1232-3C (*mata1-50 rme1-2 leu2 ura3 trp1 his4 can1*). The parents of the *a/α2⁻* diploids were AM1228-15C, AM1227-1A (*mata2-182 RME1 leu2 ura3 trp1 his4 can1* [pJM9]) and AM1227-1B (*mata2-182 rme1-2 leu2 ura3 trp1 his4 can1* [pJM9]). The *mata1-50* and *mata2-182* alleles are *Xho*I linker insertion mutations³⁵ provided by P. Siliciano. Plasmid pJM9 is YCp50 (C. Mann, personal communication) containing a *Hind*III fragment with the *MATα* allele³⁵, constructed by J. Margolske. Diploids were either selected as prototrophs or isolated by micromanipulation. The plasmid pJM9 was cured when necessary by isolating *Ura*⁻ mitotic segregants after growth in non-selective media. The *MAT* genotype of each diploid was confirmed by bioassay of mating pheromones^{36,37}. *a/α* diploids produce no pheromone, *a1⁻/α* diploids produce *α* pheromone, and *a/α2⁻* diploids produce *a* pheromone.

greater extent in *a/α* cells than is the *RME1* transcript by rehybridizing the blot displayed in Fig. 3 with an *HO* probe (probe Ylp5-BH2, ref. 9; data not shown).

We next asked whether the repression of the *RME1* transcript, observed in *a/α* diploids, depends on the *MAT* gene products that confer the ability to enter meiosis, *a1* and *α2*. Indeed, diploids with mutations at either *MATa1* or *MATα2* display the same high levels of the *RME1* transcript as the *a* and *α* haploids (Fig. 3, lanes 4, 6). The *MATα1* gene product has no role in initiation of meiosis^{11,18} and is completely dispensable for *a/α* repression of the *RME1* transcript (Fig. 3, lane 5). This analysis demonstrates both that *RME1* activity is controlled by the *MAT* gene products *a1* and *α2* and that this regulation is exerted on the *RME1* transcript level.

Effects of altered *RME1* expression

The proposed role of *RME1* activity is to inhibit meiosis⁶. This view holds that the *MAT* locus ultimately governs the sporulation competence of the cell only by regulation of *RME1* expression. Therefore, we sought to find whether the ability to enter meiosis and sporulate is dictated by *RME1* activity when that level of activity is manipulated so that it no longer reflects the mating type of the cell.

First, cells that lack the *RME1* function should be able to sporulate regardless of their *MAT* genotype. This property is one of the defining features of the previously studied *rme1* mutations^{5,6,12}. We have determined that the *rme1-2* insertion allele, a known null mutation, shares this property (Table 1). As noted above, *a1⁻/α* and *a/α2⁻* diploids are incapable of sporulation if they possess *RME1* function (*RME1/rme1-2* strains). However, such cells produce an abundance of ascospores in sporulation medium if *RME1* function is eliminated, as in the *rme1-2/rme1-2* diploids. Therefore, the *RME1* product functions in *a1⁻/α* and *a/α2⁻* cells to inhibit sporulation.

Second, sporulation should be blocked in *a/α* diploids if they overexpress the *RME1* product. We reasoned that such overproduction should result from an increased *RME1* gene dosage

because a/α repression of a single genomic copy of *RME1* does not completely abolish its expression (Fig. 3). Indeed, an a/α diploid carrying *RME1* DNA inserted into the multi-copy plasmid YEp24 (ref. 19) has 100-fold increased levels of the *RME1* transcript (data not shown). This strain sporulated <10% as well as isogenic strains that carry either the YEp24 vector alone or a YEp24-borne *rme1-2* insertion allele (Table 2). Moreover, at least 90% of the asci from the *RME1* overproducer derive from mitotic segregants that have lost the plasmid because 22 out of 23 such asci contained no plasmid-bearing spores. Thus, elevated *RME1* expression in a/α diploids inhibits ascospore production at least 100-fold.

The above experiments examined production of ascospores. We have also investigated the effects of the presence or absence of the *RME1* product on an early event in meiosis, the enhanced production of intragenic recombinants²⁰. These studies used the two alleles *his4-519* and *his4-712* (ref. 21); the transfer of a *MATa/MATα his4-519/his4-712* diploid to sporulation medium results in >300-fold stimulation in the rate of production of *His*⁺ intragenic recombinants (Table 2). This stimulation was reduced 30-fold in a/α cells carrying the multi-copy *RME1* plasmid (pAM232) but was unaffected by the multi-copy *rme1-2* plasmid (pAM238L). Similarly, $a1^-/\alpha$ cells homozygous for the *rme1-2* null allele displayed elevated intragenic recombination rates in sporulation medium; the presence of a wild-type *RME1* allele blocked this response (data not shown), in agreement with an earlier study of an *rme1* mutation²². These observations support the idea^{5,22} that the target of *RME1* control is entry into meiosis, rather than completion of sporulation.

In wild-type a/α diploids, meiosis and sporulation are blocked by the presence of either glucose or ammonia³. These nutrients also inhibit sporulation in diploids homozygous for the *rme1-2* insertion. Therefore, the *RME1* product does not mediate nutritional control of sporulation.

Discussion

We have demonstrated that the *RME1* product is an inhibitor of meiosis and sporulation and that cells specialized to undergo meiosis, a/α cells, achieve this state by repression of the *RME1* gene. Thus, the *MAT* locus governs the two fundamental processes of the yeast life cycle involved with changes in ploidy (mating and meiosis) via regulation of the transcript levels of its target genes.

The *RME1* product was originally thought to be an inhibitor of meiosis because the *rme1-1* allele is recessive in enabling a/α

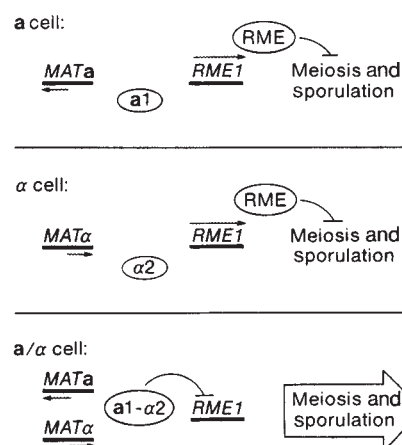


Fig. 4 Regulatory circuit linking *MAT*, *RME1* and meiosis and sporulation. To the left is drawn the mating type locus of *a*, α and a/α cells. In *a* cells, the *a1* product (encoded by *MATa*) has no known role^{5,11}; in α cells, the *α2* product (encoded by *MATα*) is a repressor of *a*-specific genes⁴¹; the *RME1* product is expressed in these cell types and blocks entry into meiosis. In a/α cells, *a1-α2* activity represses *RME1* and thereby allows meiosis and sporulation to proceed.

and α/α cells to sporulate⁶. This mutation was discovered as a genetic difference among laboratory strains and thus its precise origin and nature are unknown^{5,6}. We have shown here that a null allele of *RME1* constructed *in vitro* (*rme1-2*) also bypasses control of sporulation by the mating type locus. Not only does absence of *RME1* product allow sporulation but its presence blocks sporulation: we have found that a/α diploids that overexpress the *RME1* product (as a consequence of increased gene dosage) are defective for meiosis and spore formation. Our results thus confirm that the role of the *RME1* product is to inhibit meiosis and sporulation.

It was further proposed⁶ that the *RME1* function would be under negative control by *a1-α2* activity because *a1* and *α2* products are necessary to promote sporulation in *RME1* strains. This model postulates that *a1-α2* acts as a negative regulator, for which there is ample precedent⁷⁻⁹. Our finding that *RME1* transcript levels are repressed in a/α cells substantiates this proposal and specifies the level at which regulation is exerted

Table 2 Elevated levels of *RME1* product in a/α cells inhibit sporulation and meiotic recombination

Plasmid	Insert	Sporulation (%)	Presence of plasmid after sporulation (ascus type)					Plasmid	Insert	% His ⁺	
			4+ : 0-	3+ : 1-	2+ : 2-	1+ : 3-	0+ : 4-			Veg	Spo
pAM232	<i>RME1</i>	2	0	0	0	1	22	pAM232	<i>RME1</i>	<0.01	0.1
YEp24	None	24	4	1	0	0	1	pAM238L	<i>rme1-2</i>	<0.01	3.8
pAM238L	<i>rme1-2</i>	55	17	3	0	0	2	None	—	<0.01	3.4

Left part of table: Plasmids listed were transformed¹³ into the a/α diploid AM1788. Two or more transformants harbouring each plasmid were purified, grown to saturation in SC-uracil medium (which selects for maintenance of the plasmids), then washed and diluted 10-fold into liquid sporulation medium. The percentage sporulation was determined after 12 days at 30°C as described in Table 1. Several of the asci produced by each diploid were dissected; presence of the plasmids was determined by scoring the plasmid-borne *URA3* gene. Diploid AM1788 is homozygous for the markers *leu2 ura3 trp1 his4 can1*. Plasmid pAM232 contains a 4-kb *PvuII/XhoI* insert which includes the *RME1* gene ligated between the *PvuII* and *SalI* sites of plasmid YEp24 (ref. 19). pAM232 was digested with *BglII* and ligated in the presence of a 3-kb *BglII* fragment containing the *LEU2* gene to construct plasmid pAM238L; its structure was confirmed by complementation of an *Escherichia coli leuB* mutant and by restriction enzyme analysis. Right part: Transformants of the a/α diploid AM-S151 carrying the plasmids listed were grown to saturation in SC-uracil medium; the untransformed parent was grown in YEPD. The cultures were washed and diluted to 10⁷ cells ml⁻¹ in liquid sporulation medium, then incubated for 9 days at 30°C to allow sporulation to occur. Samples were removed before (Veg) or after (Spo) this incubation, sonicated and used to determine the frequencies of His⁺ recombinants. Total cell number was assayed by plating on YEPD and the number of His⁺ recombinants was assayed by plating on SC-histidine. For the experiments using plasmid-bearing diploids, only Ura⁺ segregants were counted; the numbers of these segregants were determined by replica-plating the YEPD and SC-histidine plates to SC-uracil. The genotype of strain AM-S151 is *MATa/MATα his4-519/his4-712 rme1-2/rme1-2 leu2/leu2 ura3/ura3 trp1/trp1 can1/+ lys1/+*.

(see Fig. 4 for summary). Haploid *a* cells express *RME1* at a high level because of the absence of the *MATa2* product; the *RME1* product prevents entry into meiosis. The same situation exists in α cells because the *MATa1* product is absent. However, in *a*/ α cells, both the *MATa1* and *MATa2* products are present, giving rise to the negative regulator *a1*- α 2. The *RME1* transcript is repressed by *a1*- α 2, and the resulting absence of the *RME1* product permits initiation of meiosis.

Although *RME1* transcript levels are greatly reduced in *a*/ α cells relative to *a* and α cells, there is clearly residual expression. This residual expression must be from the *RME1* gene because no hybridizing species is observed in *rme1-2* insertion mutants. The repression ratio for *RME1* is thus 10–20-fold. In contrast, repression of *MATa1* and *HO* by *a1*- α 2 is greater than 100-fold^{7–9}. To understand why, in a mechanistic sense, the repression of *RME1* is so much less than that of these other genes will require an understanding of the basis for repression of *RME1* by *a1*- α 2. *a1*- α 2 might act directly on *RME1* DNA itself or it might turn off synthesis of an activator of *RME1*. The observation of residual *RME1* transcript in *a*/ α cells raises the question of whether this low level has any functional significance. For example, one might imagine that a low level of *RME1* activity stimulates sporulation. We have observed, however, that sporulation of *a*/ α cells is unaffected by complete absence of *RME1* activity (due to the *rme1-2* insertion; Table 1 and A.P.M., unpublished results). The residual transcript may thus have no functional significance.

The *rme1-2* insertion mutation does not completely bypass *MAT* control of sporulation: *a1*[−]/ α and *a*/ α 2[−] strains homozygous for *rme1-2* do not sporulate as rapidly or efficiently as *a*/ α cells (Table 1 and A.P.M., unpublished results). We consider it unlikely that the *rme1-2* mutant product retains some residual activity because no *rme1-2* transcript is detectable in RNA blots (Fig. 2) and because an increased gene dosage of the *rme1-2* allele does not inhibit sporulation in *mata1*/*MATa* *rme1-2*/*rme1-2* diploids (A.P.M., unpublished results). Instead it seems likely that the mating-type locus does more to promote efficient sporulation than simply to repress *RME1* expression. For example, *a1*- α 2 could repress a second inhibitor of sporulation.

Are there genetic requirements that govern entry into meiosis in addition to the presence of *a1*- α 2 activity? For example, is diploidy required? Diploidy itself does not allow cells to sporulate^{3,4}, as noted above. Furthermore, diploidy is not even

necessary to initiate meiosis: haploids can engage in pre-meiotic DNA synthesis, recombination and even later steps in spore formation if they either express *a1*- α 2 or lack *RME1* activity (refs 3, 5, 22–24 and A.P.M., unpublished results). In contrast, diploidy is essential for completion of a normal meiosis: haploids altered such that they express *a1*- α 2 are unable to complete the meiotic divisions successfully because they lack homologous chromosomes^{23,24}. Because *a1*- α 2 activity depends on products from both alleles of *MAT*, the *MAT* locus behaves as a natural monitor of ploidy in yeast, allowing a distinction between *a* or α haploids and *a*/ α diploids. Expression of *RME1* could be a protective measure to ensure that haploids never enter what would be an aberrant meiosis.

There are two known signals required for cells to enter meiosis: one is under *MAT* control, the other depends on starvation³. Repression or inactivation of *RME1* relieves *MAT* control but not nutritional regulation. We have also found that the *cyr1-1* mutation, which bypasses nutritional control of sporulation²⁵, does not abolish the *MAT* requirement (A.P.M., unpublished results). These observations suggest that meiosis is controlled by two independent regulatory circuits. Whether there is a single gene product in *S. cerevisiae*, such as the *pat1* or *ran1* product of *Schizosaccharomyces pombe*^{26–28}, whose loss abolishes both mating type and nutritional control of sporulation, is presently unknown.

There are various possibilities for the mechanism of *RME1* action. The existence of mutations in many different genes that specifically block sporulation^{3,29} suggests that many gene products are used exclusively during sporulation or that vegetative products must assume new roles in this process. Several recent studies have documented sporulation-specific changes in cellular proteins and mRNA species^{30–34}. Perhaps the *RME1* product is a repressor of some of these genes. Another possibility is that the *RME1* product participates in covalent modification or degradation of a protein target. We anticipate that such targets will identify important early steps in meiosis.

We thank Flora Banuett, Stan Fields, Sandy Johnson, Susan Michaelis and Paul Sternberg for comments on the manuscript, and all members of our laboratory for discussions and encouragement. We also thank Paul Siliciano and Kelly Tatchell for the contribution of strains. This work was supported by NIH research grant AI 18738 (to I.H.) and by a postdoctoral fellowship (to A.M.) from Damon Runyon-Walter Winchell Cancer Fund.

Received 7 October; accepted 20 December 1985.

- Herskowitz, I. & Oshima, Y. in *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance* (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 181–209 (Cold Spring Harbor Laboratory, New York, 1981).
- Sprague, G. F. Jr, Blair, L. C. & Thorner, J. A. *Rev. Microbiol.* **37**, 623–660 (1983).
- Esposito, R. E. & Klapholz, S. in *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance* (eds Strathern, J. N., Jones, E. W. & Broach, J. R.) 211–287 (Cold Spring Harbor Laboratory, New York, 1981).
- Roman, H., Phillips, M. M. & Sands, S. M. *Genetics* **40**, 546–561 (1955).
- Kassir, Y. & Simchen, G. *Genetics* **82**, 187–206 (1976).
- Rine, J. D., Sprague, G. F. Jr & Herskowitz, I. *Molec. cell. Biol.* **1**, 958–960 (1981).
- Klar, A. J. S., Strathern, J. N., Broach, J. R. & Hicks, J. B. *Nature* **289**, 239–244 (1981).
- Nasmyth, K. A., Tatchell, K., Hall, B. D., Astell, C. & Smith, M. *Nature* **289**, 244–250 (1981).
- Jensen, R., Sprague, G. F. Jr & Herskowitz, I. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3035–3039 (1983).
- Elder, R. T., St John, T. P., Stinchcomb, D. T. & Davis, R. W. *Cold Spring Harb. Symp. quant. Biol.* **45**, 581–584 (1981).
- Strathern, J. N., Hicks, J. B. & Herskowitz, I. *J. molec. Biol.* **147**, 357–372 (1981).
- Hopper, A. K. & Hall, B. D. *Genetics* **80**, 41–59 (1975).
- Hinnen, A., Hicks, J. B. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1929–1933 (1978).
- Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354–6358 (1981).
- Rothstein, R. J. *Meth. Enzym.* **101**, 202–211 (1983).
- Southern, E. J. *molec. Biol.* **98**, 503–517 (1975).
- Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
- MacKay, V. & Manney, T. R. *Genetics* **76**, 273–288 (1974).
- Botstein, D. et al. *Gene* **8**, 17–24 (1979).
- Sherman, F. & Roman, H. *Genetics* **48**, 255–261 (1963).
- Culbertson, M. R., Charnas, L., Johnson, M. T. & Fink, G. R. *Genetics* **86**, 745–764 (1977).
- Hopper, A. K. & Hall, B. D. *Genetics* **80**, 61–76 (1975).
- Roth, R. & Fogel, S. *Molec. gen. Genet.* **112**, 295–305 (1971).
- Wagstaff, J. E., Klapholz, S. & Esposito, R. E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2986–2990 (1982).
- Matsumoto, K., Uno, I. & Ishikawa, T. *Cell* **32**, 417–423 (1983).
- Iino, Y. & Yamamoto, M. *Molec. gen. Genet.* **198**, 416–421 (1985).
- Nurse, P. *Molec. gen. Genet.* **198**, 497–502 (1985).
- Iino, Y. & Yamamoto, M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2447–2451 (1985).
- Tsuboi, M. *Molec. gen. Genet.* **182**, 1–6 (1983).
- Kraig, E. & Haber, J. E. *J. Bact.* **144**, 1098–1112 (1980).
- Clancy, M. J. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3000–3004 (1983).
- Percival-Smith, A. & Segall, J. *Molec. cell. Biol.* **4**, 142–150 (1984).
- Weir-Thompson, E. M. & Dawes, I. W. *Molec. cell. Biol.* **4**, 695–702 (1984).
- Kurtz, S. & Lindquist, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7323–7327 (1984).
- Tatchell, K., Nasmyth, K. A., Hall, B. D., Astell, C. & Smith, M. *Cell* **27**, 25–35 (1981).
- Sprague, G. F. Jr & Herskowitz, I. *J. molec. Biol.* **153**, 305–321 (1981).
- Wilson, K. L. & Herskowitz, I. *Molec. cell. Biol.* **4**, 2420–2427 (1984).
- Bolivar, F. et al. *Gene* **2**, 95–113 (1977).
- Struhl, K., Stinchcomb, D. T., Scherer, S. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1035–1039 (1979).
- Norlander, J., Kempe, T. & Messing, J. *Gene* **26**, 101–106 (1983).
- Johnson, A. D. & Herskowitz, I. *Cell* **42**, 237–247 (1985).

A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences

Dionisio Martin-Zanca*, Stephen H. Hughes† & Mariano Barbacid*

* Developmental Oncology Section and † Gene Expression Section, Basic Research Program, Frederick Cancer Research Facility, Frederick, Maryland 21701, USA

A biologically active complementary DNA clone of a transforming gene present in a human colon carcinoma contains gene sequences of both tropomyosin and a previously unknown protein tyrosine kinase. The predicted protein (641 amino acids) encoded by this oncogene seems to have been formed by a somatic rearrangement that replaced the extracellular domain of a putative transmembrane receptor by the first 221 amino acids of a non-muscle tropomyosin molecule.

TRANSFECTION of NIH 3T3 cells with genomic DNAs has allowed the identification of transforming genes in human neoplasms and in carcinogen-induced animal tumours (for recent reviews, see refs 1–3). Most of these oncogenes are members of the *ras* gene family. However, other transforming genes without apparent sequence homology to *ras* genes have been identified: those of human origin include *oncD* (carcinoma of the ascending colon⁴), *met* (MNNG-HOS cells⁵), *mel* (NK14 melanoma cells⁶), *dbl* (diffuse B-cell lymphoma⁷), and a transforming allele of *c-raf-1* (carcinoma of the stomach⁸). Two additional oncogenes, designated *mcf-3* (ref. 9) and *ret*¹⁰, became activated during transfection of NIH 3T3 cells with human DNA. Finally, a series of non-*ras* transforming genes which exhibit differentiation-stage specificity have been consistently identified by Cooper *et al.*¹¹ in a variety of human tumours.

During gene transfer assays aimed at detecting transforming genes in human tumours, our laboratory identified four different transforming genes (designated *oncA* to *oncD*) based on their distinct pattern of *Alu*-positive restriction endonuclease fragments⁴. Three of these oncogenes (*oncA*, *B* and *C*) were identified as transforming alleles of the cellular *Ha-ras-1*, *Ki-ras-2* and *N-ras* loci; however, *oncD* did not show detectable sequence homology with members of the *ras* gene family even under relaxed hybridization conditions^{4,12}. *oncD* exhibits a genetic complexity of at least 30 kilobases (kb), as all tertiary NIH 3T3 transformants analysed contain two *Alu*-positive *EcoRI* DNA fragments of 25 and 4 kb (see Fig. 2 of ref. 4). Here, we describe the isolation of a biologically active complementary DNA clone of *oncD* and its complete nucleotide sequence analysis. Information derived from these studies has allowed us to determine that *oncD* was generated by a somatic rearrangement that resulted in the juxtaposition of two truncated loci, one of which codes for a non-muscle tropomyosin and the other for a putative tyrosine-specific protein kinase.

Molecular cloning of *oncD* cDNA

Three overlapping λ recombinant clones representing about 20 kb of *oncD* were isolated from a genomic library prepared from 106-63 cells, a third-cycle NIH 3T3 transformant derived from *oncD* (D.M.-Z., B. Cabrer and M.B., unpublished). Several *Alu*-negative DNA fragments were subcloned in vector pUC18 and used as probes to identify *oncD*-specific transcripts in NIH 3T3 transformants. Two of these subclones, pDM6, which contained a 1.4-kb *BamHI*–*EcoRI* insert, and pDM8, which carried a 2.7-kb *KpnI* insert, identified a 2.5-kb poly(A)-containing messenger RNA in *oncD*-derived transformants. Significant levels of mRNA of this size were also found in normal NIH 3T3 cells as well as in *ras*-derived transformants (data not shown).

A cDNA library consisting of 10⁶ clones was prepared as described by Helfman *et al.*¹³. DH5 *Escherichia coli* cells were transformed with 25 ng of cDNA prepared by reverse transcription of polyadenylated mRNA isolated from 106-63 cells and ligated to pUC18 and pUC19 vectors using *SalI* and *EcoRI* linkers. Nitrocellulose filters containing these colonies were hybridized with the 1.4-kb *BamHI*–*EcoRI* insert of pDM6 and with the 2.7-kb *KpnI* insert of pDM8 as described previously¹⁴. About 200 hybridizing colonies were selected for further analysis.

Presence of non-muscle tropomyosin sequences

The existence of *oncD*-related polyadenylated mRNA in untransformed NIH 3T3 cells suggested that a significant fraction of the *oncD*-positive cDNA clones might represent normal mouse cellular transcripts. Partial nucleotide sequence analysis of the pDM6 and pDM8 inserts was used to identify a domain that did not hybridize with NIH 3T3 mRNAs. Comparison of pDM6 and pDM8 *oncD* sequences with those stored in gene databanks revealed that a stretch of 204 base pairs (bp) of pDM6 and 66 bp of pDM8 were almost identical (>96% homology) to a clone (hTM_{NM-1}) of a human non-muscle tropomyosin pseudogene characterized previously¹⁵ (Fig. 1). Moreover, these regions had open reading frames whose deduced amino-acid sequences also exhibited striking homology (over 90%, taking into account conserved amino-acid substitutions) to the amino terminus (residues 1–43 in the case of pDM6) and to an internal domain (residues 177–197 in the case of pDM8) of a tropomyosin molecule isolated from horse platelets¹⁶ (Fig. 1). Finally, pDM6 and pDM8 sequences located at the 3' boundaries of their respective regions of homology to hTM_{NM-1}, closely matched those of consensus donor splicing sites (not shown). These results suggested strongly that pDM6 and pDM8 contained the first coding exon and an internal exon, respectively, of a non-muscle tropomyosin gene. Tropomyosins are known to be expressed at high levels in most cell types, explaining the existence of abundant *oncD*-related transcripts in untransformed NIH 3T3 cells.

The 204-bp region of pDM6 that is homologous to hTM_{NM-1} includes 70 bp of 5'-untranslated sequence⁵ (Fig. 1). Noncoding sequences evolve faster than their translated counterparts. Thus, we reasoned that this 70-bp segment of pDM6 might have diverged sufficiently from its murine counterpart to allow identification of cDNA clones carrying *oncD* sequences from those containing mouse tropomyosin transcripts. The 70-bp 5'-untranslated sequences were isolated as part of a *BamHI*–*NcoI* DNA fragment of pDM6 which extended from the 5' *BamHI* boundary of the insert to the convenient *NcoI* site which contains the putative initiator codon (ATG) of the tropomyosin gene

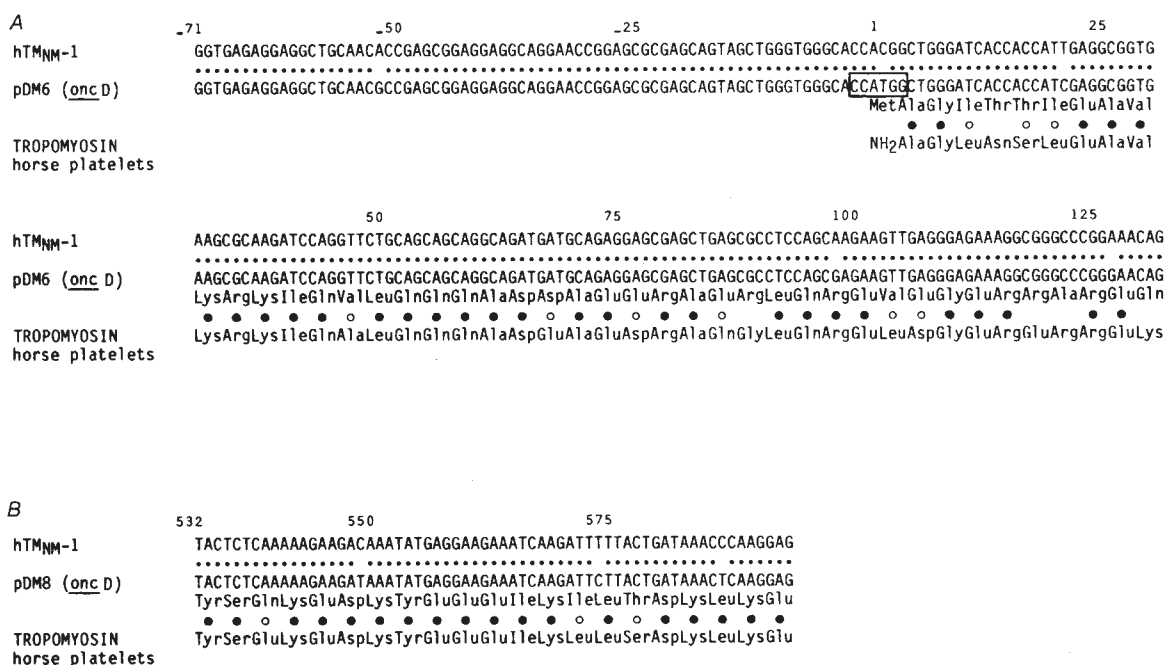


Fig. 1 Nucleotide sequence homology between pDM6 (**A**) and pDM8 (**B**) subclones of *oncD* and hTM_{NM}-1, a clone of a human non-muscle tropomyosin pseudogene¹⁵. Dots indicate base pairs that are identical in the two sequences. Numbers correspond to numbering of nucleotides in the hTM_{NM}-1 sequence, with +1 as the deoxyadenosine residue of the putative initiator codon (ATG) which has been replaced by a non-functional ACG triplet in hTM_{NM}-1. The box indicates the convenient *Nco*I site used in generating the 209-bp *Bam*HI-*Nco*I probe used to identify *oncD*-specific cDNA clones (see text). The amino-acid sequences deduced from the open reading frames of pDM6 (nucleotides 1-132) and pDM8 are compared with the sequence of a tropomyosin molecule isolated from horse platelets¹⁶. ●, ○, Identical and conserved amino-acid residues, respectively.

(Fig. 1). Of the 200 positive cDNA clones, only 11 specifically hybridized with this probe; these colonies were picked, grown up and their plasmids isolated by standard techniques¹⁷. One of the plasmids, pDM10-1, contained a 2.3-kb *Sal*I-*Eco*RI DNA insert which is approximately the size of the mRNAs identified in 106-63 cells by *oncD* specific probes.

Nucleotide sequence of *oncD* cDNA

The complete nucleotide sequence of the 2,301-bp *Sal*I-*Eco*RI insert of pDM10-1 was established by the method of Maxam and Gilbert¹⁸ (see Fig. 2). pDM10-1 contains a 1,923-bp open reading frame capable of directing the synthesis of a polypeptide of 641 amino acids. The open reading frame is flanked by 232 bp of 5'-untranslated sequence and 143 bp of 3'-noncoding sequence in which no putative polyadenylation signal could be identified. It is possible that pDM10-1 does not contain all of the untranslated sequences of *oncD*. However, based on the size of the mRNA, only 100-200 bp are absent.

The most relevant features of this sequence are: (1) nucleotides -232 to -64 (position +1 corresponds to the deoxyadenosine of the initiator ATG codon) represent a stretch of G + A-rich sequences in which the sequence GGAGGAGCA is repeated 12 times (with one imperfection in two cases). This repeat is interrupted by the hexanucleotide GGA^CCA on three occasions and by repeated GGA triplets on two occasions. A similar arrangement of repetitive sequences (a GGGGAGGA sequence interrupted by the hexanucleotide GCAGGA) exists in the simple IR3 (internal repeat 3) repeat array of Epstein-Barr virus¹⁹. The biological significance of this domain of the *onc* cDNA is being investigated. (2) Sequences located between positions -63 and +664 correspond to the first seven coding exons of a non-muscle tropomyosin gene, as determined by comparison of these sequences with those of genomic clones of the tropomyosin gene (A. MacLeod, personal communication)

and of *oncD* (our unpublished observations). The boundaries between each of the exons are indicated by vertical arrowheads in Fig. 2. Nucleotides -63 to -1 correspond to the sequences used to identify pDM10-1 and represent untranslated sequences of the first coding exon of the tropomyosin locus present in *oncD*. (3) The homology between pDM10-1 and genomic tropomyosin sequences ends abruptly at nucleotide +664, which corresponds to the last base of the tropomyosin exon 7. This result suggests that *oncD* is a hybrid gene generated by a genetic rearrangement that disrupted a non-muscle tropomyosin gene between coding exons 7 and 8. (4) Sequences located between nucleotides +1,015 and +1,773 share extensive homology with retroviral oncogenes coding for tyrosine-specific protein kinases such as *v-abl*, *v-erb-B*, *v-fes/fps*, *v-fgr*, *v-fms*, *v-ros*, *v-src* and *v-yes*²⁰. By introducing appropriate gaps, the homology between the sequences coding for the catalytic domains of these retroviral protein tyrosine kinases and the *oncD* sequence was as high as 51%. Similar results have been obtained with cellular genes known to possess tyrosine-specific protein kinase activity, for example those encoding the insulin²¹ and epidermal growth factor (EGF)²² receptors. Finally, a lower degree of homology was observed when these *oncD* sequences were compared with retroviral oncogenes whose products exhibit serine/threonine-specific protein kinase activity, for example, *v-mos* and *v-raf*²⁰. Taken together, these results strongly suggest that the carboxy-terminal moiety of *oncD* was derived from a locus that codes for a tyrosine-specific protein kinase.

Protein tyrosine kinase gene

Figure 3 illustrates the striking homology between the carboxy-terminal moiety of the putative product of *oncD* and the common catalytic domain of retroviral and cellular tyrosine-specific protein kinases. The putative *oncD* product has 63 of the 70 amino-acid residues common to all known protein tyrosine kinases,

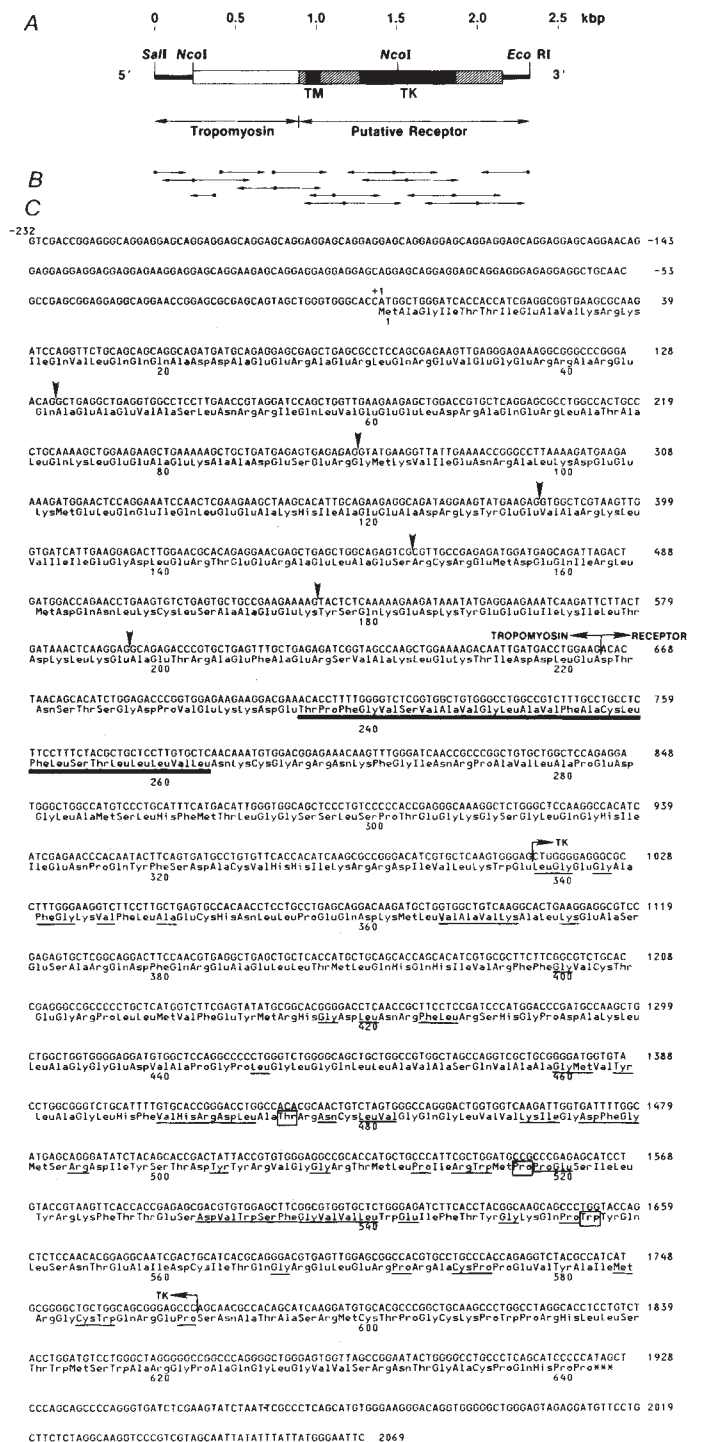
Fig. 2 Molecular characterization of a cDNA clone of *oncD*. **A**, Schematic diagram of clone pDM10-1, showing the *Sal*I and *Eco*RI cleavage sites that delimit the boundaries of the pDM10-1 insert, and the *Nco*I cleavage sites, one of which (near the 5' end) contains the initiator ATG codon. The solid line indicates untranslated sequences; the large box represents coding sequences, in which those derived from tropomyosin ($\square$) and the putative receptor gene ($\blacksquare$) are indicated. The transmembrane domain (TM) and the region exhibiting sequence homology with tyrosine-specific protein kinases (TK) are superimposed. **B**, Nucleotide sequencing strategy. **C**, Nucleotide sequence of the 2,301-bp *Sal*I-*Eco*RI insert of pDM10-1 and the deduced amino-acid sequence. Nucleotides are numbered from the deoxyadenosine residue of the initiator ATG codon and amino-acid residues from the presumed initiating methionine. The solid black underlining indicates the putative transmembrane domain. Amino-acid residues shared by other tyrosine-specific protein kinases are indicated by thin underlining. Boxed residues are those shared by other protein tyrosine kinases, but not by *oncD*. Sequences derived from the tropomyosin gene are separated from the putative receptor locus by horizontal arrows. Vertical arrowheads indicate the limits of the seven tropomyosin coding exons.

including 28 of the 31 residues shared between these protein tyrosine kinases and, the prototype serine/threonine-specific protein kinases²³. The conserved residues include Tyr 503 (Tyr 416 in pp60^{v-src}), which lies within the catalytic domain and is the target of autophosphorylation reactions, and Lys 367 (Lys 295 in pp60^{v-src}), which forms part of the ATP-binding site. *oncD* Lys 367 is located 21 residues from the glycine-rich sequence Gly-Glu-Gly-Ala-Phe-Gly, as in other protein kinases²³. Other highly conserved residues shared by known protein kinases and believed to be important for catalysis, for example, Arg-Asp-Leu (residues 472-474) and Asp-Phe-Gly (residues 491-493), are also present in the putative *oncD* protein (Fig. 3).

However, the predicted *oncD* product exhibits some unique features. Three residues common to all known protein tyrosine kinases, that is, Ala 389, Ala 430 and Tyr 463 in pp60^{v-src} (ref. 22), have been replaced by Thr 476, Pro 518 and Trp 551, respectively, in *oncD*. Interestingly, the substitution of Ala 430 with valine drastically decreases the protein tyrosine kinase activity of pp60^{v-src} (ref. 24). In addition, the putative *oncD* protein contains five additional residues, Leu-Ala-Val-Ala-Ser (positions 450–454, Fig. 3), in the centre of the catalytic domain. These observations suggests that *oncD* contains a novel protein tyrosine kinase gene.

The homology of the putative *oncD* product also extends to prototype cellular protein tyrosine kinases such as the EGF²² and insulin²¹ receptors (Fig. 3). Analysis of amino-acid sequences in domains other than the amino-terminal tropomyosin moiety (amino-acid residues 1–221) and the catalytic protein tyrosine kinase domain (amino-acid residues 339–591) revealed the presence of a long stretch of 25 hydrophobic amino acids (residues 237–262) surrounded by basic residues, a feature highly indicative of a transmembrane domain (Figs 2,3). Thus, the non-tropomyosin moiety of *oncD* might have been contributed by a gene coding for a growth factor receptor with protein tyrosine kinase activity.

Direct comparison of *oncD* sequences outside the tropomyosin domain with those of the genes coding for the EGF²² and insulin²¹ receptors clearly established that neither of these two loci was involved in the generation of *oncD*. A similar statement could be made for the *c-abl* locus whose partial sequence has already been published²⁵. However, this is not the case for the human homologues of other retroviral protein tyrosine kinase oncogenes with which *oncD* shares extensive sequence homology. To determine whether *oncD* contains a truncated homologue of a known retroviral oncogene or a novel protein tyrosine kinase locus, human DNA was hybridized with probes specific for retroviral oncogenes known to code for tyrosine-specific protein kinases, including *v-fes*, *v-fgr*, *v-fms*, *v-ros*, *v-src*



and v-yes²⁰ (Fig. 4). The resulting DNA fragments were compared with those obtained using a probe specific for the receptor domain of *oncD*. Figure 4 shows a representative experiment. The pattern of *Pst*I or *Pvu*II fragments characteristic of the *oncD* receptor domain was clearly distinct from those obtained with the retroviral probes. These hybridization experiments, together with the amino-acid sequence comparisons discussed above, clearly establish that the putative protein tyrosine kinase receptor gene involved in the generation of *oncD* represents a novel human locus. Whether this gene codes for the receptor of a known growth factor is an interesting possibility that must await identification and biochemical characterization of the normal gene product.

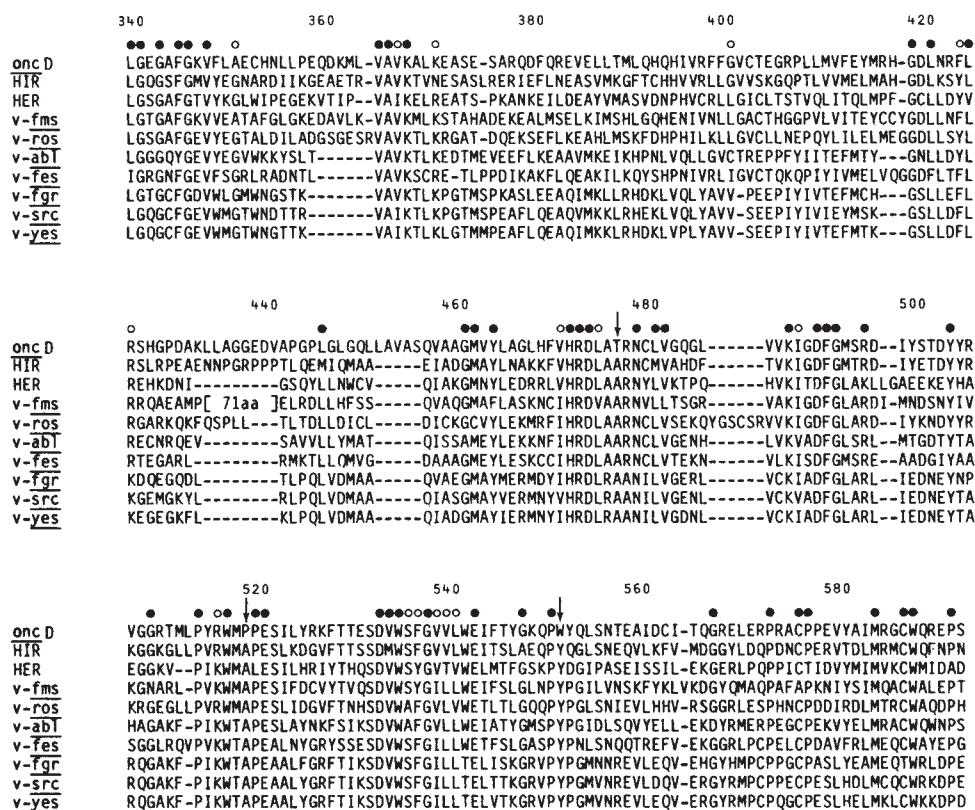


Fig. 3 Amino-acid homology of the putative *oncD* protein with the catalytic domain of tyrosine-specific protein kinases. Amino-acid residues 339-592 of the *oncD* protein were aligned with those corresponding to the homologous regions of the human insulin receptor (HIR), the human EGF receptor (HER) and the retroviral oncogenes *v-fms*, *v-ros*, *v-abl*, *v-fes*, *v-fgr*, *v-src* and *v-yes*. ●, *oncD* residues shared by at least eight of the nine protein tyrosine kinases listed; ○, conserved amino-acid substitutions. Vertical arrows indicate those residues shared by all other protein tyrosine kinases, but not by *oncD*.

Transforming activity of *oncD* cDNA

To establish whether pDM10-1 contains a functional cDNA clone of *oncD*, we investigated its potential transforming activity in gene transfer assays. For this purpose we constructed two plasmids, designated pDM14 and pDM16, in which the coding sequences of pDM10-1 were linked to a human metallothionein (IIA) promoter (derived from p84H²⁶, given by Michael Karin) and to a Moloney murine sarcoma virus (M-MSV) long terminal repeat (LTR) (derived from pm1SP²⁷, given by G. Vande Woude) (Fig. 5). In addition, pDM14 and pDM16 contained a polyadenylation signal derived from simian virus 40 (SV40)²⁸. As shown in Fig. 5, pDM16 induced the appearance of 2×10^4 foci per μg of insert DNA when added to NIH 3T3 cells in gene transfer assays²⁹, a transforming activity comparable to that of DNA clones of acute transforming retroviruses and cellular *ras* oncogenes; the transforming activity of pDM14 (6×10^3 focus-forming units (FFU) per μg) was somewhat lower. Whether this result reflects the more limited promoting activity of the metallothionein IIA regulatory sequences or is due to the fact that this construct does not contain any of the 232 bp of 5'-untranslated sequence of pDM10-1, remains to be determined. In this regard, it is interesting to note that pDM10-1 is capable of inducing low levels of NIH 3T3 transformation without the addition of a eukaryotic promoter (Fig. 5). We are presently investigating whether the G/A-rich domain of pDM10-1 has low levels of promoter activity. In any case, the results summarized in Fig. 5 establish that pDM10-1 contains a functional cDNA clone of *oncD*.

A somatic rearrangement generated *oncD*

As *oncD* was identified by transfecting human tumour DNA into NIH 3T3 cells, it was important to establish whether the genetic rearrangement responsible for its malignant activation occurred during experimental manipulations or was already present in the original colon carcinoma (tumour 2033). Two *oncD* probes capable of recognizing DNA fragments encompassing the breakpoints of the two loci present in *oncD* were

generated. As shown in Fig. 6A, the tropomyosin-specific probe (2.7-kb *KpnI* insert of pDM8) should detect an 11-kb *EcoRI* DNA fragment in normal human DNA and a 23-kb *EcoRI* fragment in DNAs containing *oncD*. Similarly, a probe derived from the kinase domain of *oncD* cDNA (1.2-kb *BalI-EcoRI* insert of pDM10-1) should recognize a normal 14-kb *BamHI* DNA fragment of human DNA and a rearranged 7-kb *BamHI* DNA fragment of *oncD* DNA (Fig. 6B). Southern transfer analysis of DNAs isolated from colon carcinoma 2033 and from first- and third-cycle NIH 3T3 transformants derived from *oncD*⁴, revealed the presence of the rearranged 23-kb *EcoRI* and 7-kb *BamHI* DNA fragments diagnostic of *oncD*. These results established that *oncD* was not activated during gene transfer but was present in the original human tumour DNA. The normal 11-kb *EcoRI* and 14-kb *BamHI* DNA fragments were also present in the original tumour DNA, demonstrating its heterozygous nature.

Finally, to test whether the rearrangement responsible for the activation of *oncD* was an inherited genetic abnormality or occurred during tumour development, DNA was isolated from normal colonic tissue adjacent to tumour 2033 (provided by S. A. Aaronson). This DNA exclusively exhibited the normal 11-kb *EcoRI* and 14-kb *BamHI* DNA fragments characteristic of normal human DNAs (Fig. 6). These results establish unequivocally that *oncD* was generated by a somatic rearrangement that brought together two independent loci during development of a colon carcinoma.

Discussion

Here we have described the generation of a human transforming gene by a genetic rearrangement that resulted in the juxtaposition of two unrelated loci. One of the two genes involved in the generation of *oncD* appears to be a novel receptor gene with associated protein tyrosine kinase activity. That a protein tyrosine kinase receptor gene could be activated as an oncogene was first illustrated by Downward *et al.*, who demonstrated that the *v-erb-B* oncogene was derived from the EGF receptor locus³⁰. It has been proposed that truncation of the amino-

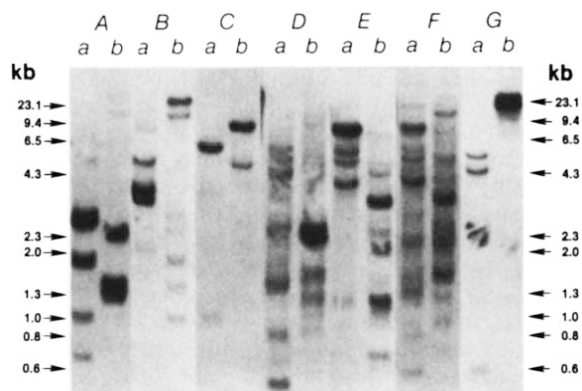


Fig. 4 A novel human locus in *oncD*. The figure shows a Southern transfer analysis of normal human DNA digested with *Pst*I (a) and *Pvu*II (b) restriction endonucleases and hybridized with probes specific for: A, *oncD* (1.2-kb *Ball*-*Eco*RI insert of pDM10-1); B, *v-fms* (409-bp *Pst*I insert of pSM7C, ATCC no. 41016); C, *v-ros* (844-bp *Eco*RI-*Pvu*II insert of p-ros, ATCC no. 41051); D, *v-fes* (459-bp *Pst*I insert of pST4, ATCC no. 41014); E, *v-fgr* (952-bp *Pvu*II-*Kpn*I insert of GR-FeSV; ref. 31); F, *v-src* (2,950-bp *Eco*RI insert of pEcoRIB, ATCC no. 41005); and G, *v-yes* (550-bp *Pst*I-*Eco*RI insert of pY-73; ref. 39).

Methods. Restricted DNAs (20 µg) were applied to 1.2% (w/v) agarose gels, electrophoresed and transferred to nitrocellulose paper as described elsewhere⁴⁰. Filters were hybridized for 48 h to 2×10^7 c.p.m. of the corresponding nick-translated probe under stringent conditions (50% (v/v) formamide, 42 °C) for the *oncD* probe and under relaxed conditions (30% (v/v) formamide, 42 °C) for the rest of the probes. Filters were exposed to Kodak XAR-5 film at -70 °C for 20 h with the help of intensifier screens. Co-electrophoresed λ *Hind*III DNA fragments served as relative molecular mass (M_r) markers.

terminal (extracellular) domain may lock the EGF receptor in its active state, thus mimicking constitutive activation by EGF²². However, it appears that an additional deletion in the carboxy-terminal end of the EGF receptor molecule is necessary for full neoplastic potential (ref. 22 and A. Ullrich, personal communication). In the case of *oncD*, it seems that a putative receptor gene has been decapitated while maintaining the transmembrane and catalytic domains without any other detectable rearrangements. However, the homology between other protein tyrosine kinases and *oncD* ends at residue 592 (Fig. 3), about 15 residues from the consensus carboxy-terminal boundary²². Identification of additional genetic alterations in the protein tyrosine kinase domain of *oncD* must await the molecular cloning of its normal allele.

v-fgr, the oncogene of the Gardner-Rasheed strain of feline sarcoma virus (GR-FeSV), also encodes a protein tyrosine kinase that carries a domain (128 amino acids) of γ -actin, another cytoskeletal protein³¹. Normal *c-fgr*, unlike the predicted *oncD* product, apparently is not a receptor^{32,33}. Moreover, the 118 amino-terminal residues of the *v-fgr* protein are contributed by viral *gag* gene sequences³¹. However, it is possible that the respective cytoskeletal domains of the *v-fgr* and *oncD* products play a part in redirecting their subcellular localization, preventing them from interacting with their normal substrates. In this regard, it is worthwhile noting that tropomyosins do not contain leader signal sequences that would be expected to direct transport through lipid bilayers. Thus, it is possible that the *oncD* protein cannot be anchored to the cellular membrane as predicted by its receptor-like structure. Generation of specific antibodies as well as appropriate deletion mutants should help to resolve the role of the cytoskeletal domains in the transforming properties of the *v-fgr* and *oncD* proteins.

Molecular characterization of genomic and cDNA clones of tropomyosin genes has indicated that some tropomyosin isoforms may be encoded by a single structural gene that used

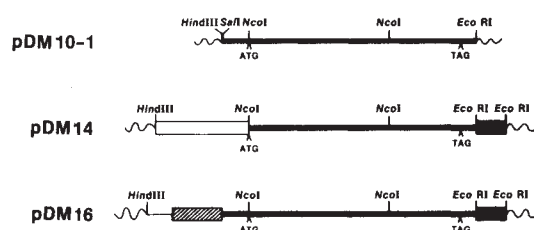


Fig. 5 Transforming activity of *oncD* cDNA. The figure shows a schematic diagram of clones pDM10-1, pDM14 and pDM16. (~), pUC18 vector sequences; —, *oncD* cDNA sequences; □, promoter sequences of the human metallothionein IIA gene; ▨, M-MSV LTR; ■, SV40 polyadenylation sequences. The restriction endonuclease sites used to construct pDM14 and pDM16 from pDM10-1 as well as the locations of the beginning (ATG) and end (TAG) of the *oncD* cDNA coding sequences are indicated. The transforming activities of pDM10-1, pDM14 and pDM16 (as supercoiled plasmids) in gene transfer assays using NIH 3T3 mouse fibroblasts as recipient cells²⁹ were, respectively, 5×10^5 and 2×10^5 FFU per µg of transfected plasmid DNA. pUC18 alone did not elicit focus formation.

Methods. pDM14 was generated by replacing the 249-bp *Hind*III-*Nco*I fragment of pDM10-1 with an 840-bp *Hind*III-*Nco*I fragment of p84H containing the human metallothionein IIA promoter²⁷. pDM16 was constructed by inserting a 750-bp *Eco*RI-*Sma*I fragment of pm1SP²⁶, containing the U3 and R regions of M-MSV and 275-bp of mink DNA 5'-flanking sequences (thin line), into the *Sal*I site of pDM10-1. The *Eco*RI and *Sal*I sites were blunt-ended before ligation. pDM14 and pDM16 also contained a 237-bp *Bcl*I-*Bam*HI DNA fragment carrying the SV40 polyadenylation signal. The *Bcl*I and *Bam*HI cleavage sites were replaced by *Eco*RI sites.

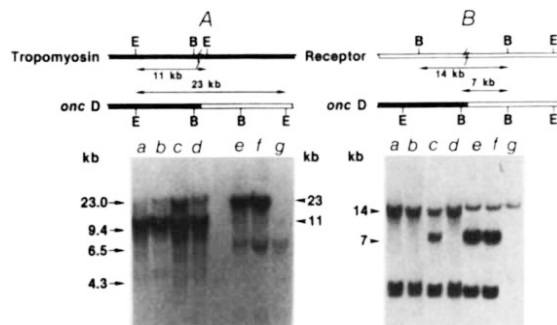


Fig. 6 *oncD* was generated by a somatic rearrangement in a human colon carcinoma. **Top**, schematic representation of DNA fragments derived from: A, *Eco*RI and B, *Bam*HI digestion of normal human DNAs and DNAs containing *oncD*. Tropomyosin sequences are represented by a solid black line whereas those of the protein tyrosine kinase gene are represented by an open line. The zig-zag line indicates the approximate location of the break-point that led to the generation of *oncD*. Only those *Eco*RI (E) and *Bam*HI (B) cleavage sites needed to define the diagnostic DNA restriction fragments are indicated. **Bottom**, Southern transfer analysis of DNAs isolated from: a, normal human lymphocytes; b, colon carcinoma 1853; c, colon carcinoma 2033, the tumour in which *oncD* was identified; d, normal colonic tissue adjacent to tumour 2033; e, first-cycle NIH 3T3 transformant derived from *oncD*; f, third-cycle NIH 3T3 transformant derived from *oncD*; g, NIH 3T3 cells. Analysis was carried out as described in Fig. 5 legend, except that 0.7% (w/v) agarose gels were used. Nitrocellulose filters were hybridized under stringent conditions to 6×10^7 c.p.m. of nick-translated probes specific for A, tropomyosin sequences (2.7-kb *Kpn*I insert of pDM8) or B, protein tyrosine kinase sequences (1.2-kb *Ball*-*Eco*RI fragment of pDM10-1). Filters were exposed to Kodak XAR-5 film at -70 °C for 20 h with the help of intensifier screens. Co-electrophoresed λ *Hind*III DNA fragments serves as M_r markers. DNA fragments which defined the genetic rearrangement that generated *oncD* are indicated by arrowheads; these include the 11-kb *Eco*RI and 7-kb *Bam*HI normal human DNA fragments and the 23-kb *Eco*RI and 14-kb *Bam*HI *oncD*-specific DNA fragments.

alternative RNA splicing mechanisms. In *Drosophila*, embryonic and thoracic tropomyosins are encoded by the same gene via alternative splicing of the last exon, which codes for the different 27 carboxy-terminal residues that distinguish the two tropomyosin isoforms³⁴. A similar mechanism appears to be responsible for the synthesis of α -tropomyosin isoforms of smooth and striated muscle of rats³⁵. Finally, alternative splicing of both internal and carboxy-terminal exons is also responsible for the synthesis of a smooth muscle-like tropomyosin expressed in human non-muscle cells and of the human skeletal muscle β -tropomyosin³⁶. Interestingly, *oncD* contains the entire tropomyosin coding sequences except for the carboxy-terminal 27 amino-acid residues presumably coded for by the missing coding exon 8 (ref. 15 and our unpublished observations). *c-abl*, the other protein tyrosine kinase oncogene implicated in human neoplasia³⁷, becomes activated by a chromosomal translocation that causes the replacement of its first coding exon with the 5' domain of another locus³⁸. Recently, Y. Ben-Neriah and D. Baltimore have shown that the normal *c-abl* locus codes for at least four different proteins by alternative splicing of the first coding exon (personal communication). It is possible that complex transcriptional processes required for the synthesis of multiple proteins by a single gene may facilitate the occurrence of

genetic rearrangements and the formation of fusion products capable of inducing malignant transformation.

Molecular characterization of novel transforming genes such as *oncD* will not only expand our current understanding of the molecular biology of neoplasia, but will prove to be a fruitful source of 'mutant' genes that should help us to elucidate the pathways involved in the control of cell growth and proliferation.

oncD to be designated *trk*

Most oncogenes have been designated according to their origin (that is, *src* for sarcoma, *ras* for rat sarcoma, *abl* for Abelson murine leukaemia virus, etc.). Although *oncD* was initially identified in a colon carcinoma⁴, apparently this oncogene is not frequently activated in this type of tumour (D.M.-Z. and M.B., unpublished observations). Therefore, we suggest that this rearranged oncogene be designated by a conventional three-letter abbreviation, *trk* (pronounced 'track'), that reflects its molecular structure: tropomyosin(t)-receptor(r)-kinase(k).

We thank Linda K. Long and Jeffrey Keefer for technical assistance, and Alexander MacLeod for communicating his results before publication. This research was sponsored by the NCI, Department of Health and Human Services, under contract no. N01-C0-23909 with Litton Bionetics, Inc.

Received 11 November 1985; accepted 8 January 1986.

- Varmus, H. E. *A. Rev. Genet.* **18**, 553-612 (1984).
- Barbacid, M. in *Important Advances in Oncology 1986* (eds Devita, V., Hellman, S. & Rosenberg, S.) (Lippincott C., Philadelphia, in the press).
- Balmain, A. *Br. J. Cancer* **51**, 1-7 (1985).
- Pulciani, S. *et al. Nature* **300**, 539-542 (1982).
- Cooper, C. S. *et al. Nature* **311**, 29-33 (1984).
- Padua, R. A., Barrass, N. & Currie, G. A. *Nature* **311**, 671-673 (1984).
- Eva, A. & Aaronson, S. A. *Nature* **316**, 273-275 (1985).
- Shimizu, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5641-5645 (1985).
- Fasano, O., Birnbaum, D., Edlund, L., Fogh, J. & Wigler, M. *Molec. cell. Biol.* **4**, 1695-1705 (1984).
- Takahashi, M., Ritz, J. & Cooper, G. M. *Cell* **42**, 581-588 (1985).
- Cooper, G. M. & Lane, M. A. *Biochim. biophys. Acta* **738**, 9-20 (1984).
- Notario, V., Sukumar, S., Santos, E. & Barbacid, M. *Cold Spring Harb. Conf. Cell Proliferation* **Cancer** **11**, 425-432 (1984).
- Helfman, D. M., Feramisco, J. R., Fiddes, J. C., Thomas, G. P. & Hughes, S. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 31-35 (1983).
- Hanahan, D. & Meselson, M. *Gene* **10**, 63-67 (1980).
- MacLeod, A. R. & Talbot, K. *J. molec. Biol.* **167**, 523-537 (1983).
- Lewis, W. G., Cote, G. P., Mak, A. S. & Smilie, L. B. *FEBS Lett.* **156**, 269-273 (1983).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual*, 90 (Cold Spring Harbor Laboratory, New York, 1982).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Baer, R. *et al. Nature* **310**, 207-211c (1984).

- Weiss, R., Teich, N., Varmus, H. & Coffin, J. (eds) *RNA Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1985).
- Ullrich, A. *et al. Nature* **313**, 756-761 (1985).
- Ullrich, A. *et al. Nature* **309**, 418-425 (1984).
- Hunter, T. & Cooper, J. A. *A. Rev. Biochem.* **54**, 897-930 (1985).
- Bryant, D. & Parsons, J. T. *J. Virol.* **45**, 1211-1216 (1983).
- Groffen, J., Heisterkamp, N., Reynolds, F. H. & Stephenson, J. R. *Nature* **304**, 167-169 (1983).
- Karin, M. *et al. Nature* **308**, 513-519 (1984).
- Blair, D. G. *et al. Science* **212**, 941-943 (1981).
- Fitzgerald, M. & Shenk, T. *Cell* **24**, 251-260 (1981).
- Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
- Downward, J. *et al. Nature* **307**, 521-526 (1984).
- Naharro, G., Robbins, K. C. & Reddy, E. P. *Science* **223**, 63-66 (1984).
- Tronick, S. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 6595-6599 (1985).
- Parker, R. C., Mardon, G., Lebo, R. V., Varmus, H. E. & Bishop, J. M. *Molec. cell. Biol.* **5**, 831-838 (1985).
- Basi, G. S., Boardman, M. & Sorti, R. *Molec. cell. Biol.* **4**, 2828-2836 (1984).
- Ruiz-Opazo, N., Weinberger, J. & Nadal-Ginard, B. *Nature* **315**, 67-70 (1985).
- MacLeod, A. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 7835-7839 (1985).
- Heisterkamp, N. *et al. Nature* **306**, 239-242 (1983).
- Shtivelman, E., Lifshitz, B., Gale, R. P. & Canaani, E. *Nature* **315**, 550-554 (1985).
- Kitamura, N., Kitamura, A., Toyoshima, K., Hirayama, Y. & Yoshida, M. *Nature* **297**, 205-208 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).

LETTERS TO NATURE

Inflation and shadow matter

Lawrence M. Krauss*, Alan H. Guth*, David N. Spergel, George B. Field & William H. Press

Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, Massachusetts 02138, USA

* Present addresses: Department of Physics, Yale University, New Haven, Connecticut 06511, USA (L.M.K.); Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA (A.H.G.).

Recent work on superstring theories has prompted interest in 'shadow matter', exotic matter which interacts only gravitationally with normal matter. Green and Schwarz^{1,2} have shown that the only anomaly-free type 1 superstring theories are those with gauge group SO_{32} . Gross *et al.*³ have constructed a superstring theory resulting in a gauge group $E_8 \times E_8$. Such a theory could result, at low energies, in the existence of two sectors: an 'observed' sector associated with all familiar particles and interactions, and another

'hidden' sector, previously discussed in various low-energy supersymmetry models (see, for example, ref. 4 and references therein). The particles of the hidden E_8 sector couple only through gravitational interactions with ordinary matter. Kolb, Seckel and Turner⁵ have explored some of the astrophysical and cosmological implications of the existence of such shadow matter. We demonstrate here that if, in the early Universe, an inflationary phase⁶⁻⁸ is associated with the breaking of one of the symmetries in the $E_8 \times E_8$ theory, this strongly constrains the physics of both sectors if shadow matter is to be the missing mass in the Universe.

We begin by assuming that inflation occurs because of symmetry breaking in the observed E_8 sector. After inflation has ended, the abundance of all observed particles is determined by particle production and reheating as the scalar field responsible for inflation oscillates and relaxes to its true minimum by particle production^{9,10}. As is well known for the case of gravitinos in low-energy supergravity models, the production of particles which are only gravitationally coupled to ordinary matter will be suppressed during reheating compared with ordinary matter, by powers of the reheating temperature over the Planck mass. This allows one to place upper bounds on the

reheating temperature after inflation, so that gravitinos are adequately suppressed¹¹⁻¹⁷.

Gravitinos can be produced by first-order gravitational interactions¹²⁻¹⁴. If the shadow matter candidates are heavier than gravitinos, we expect that the particles in the hidden sector could be produced only by processes involving virtual intermediate gravitons or gravitinos (and thus an extra power of the inverse Planck mass, M_{Pl}^{-1}). We consider this case first. The lowest-order cross-sections, if particle masses are small compared with thermal energies, are of the order of

$$\sigma \approx N_s T^2 / M_{\text{Pl}}^4 \quad (1)$$

where N_s gives the number of final helicity states of shadow matter which can be produced at temperature T . (Such processes were also considered in the context of gravitinos by Ellis *et al.*¹⁸.) Each particle in the gas can scatter only with the particle with correct quantum numbers to allow production of the intermediate-state graviton or gravitino. This leads to a production rate per particle in the gas at temperature T

$$\Gamma(T) = n_T \langle \sigma v \rangle \approx N_s T^5 / M_{\text{Pl}}^4 \quad (2)$$

where $n_T \approx T^3$ is the number density of each particle species in the radiation gas. (Note that a greater number of scattering partners are accessible if one considers processes of higher order in other couplings. At best this would yield rates which are multiplied by an additional factor $N_s \alpha$, where N_s is the number of helicity states in the radiation gas at temperature T , and α is the relevant fine-structure constant. These two factors will tend to cancel, yielding rates of the order of those given in equation (2).)

The production of shadow matter begins after reheating at temperature T_R , and continues at a rate near $\Gamma(T_R)$ for a time given by the inverse of the Hubble expansion rate, $H \sim N_s^{1/2} T_R^2 / M_{\text{Pl}}$. The ratio r of the number density of shadow matter particles produced to the number density of particles in the radiation gas is then given approximately by

$$\Gamma \approx (N_s / N_s^{1/2}) T_R^3 / M_{\text{Pl}}^3 \quad (3)$$

Finally, this ratio is approximately equal to the ratio of shadow matter particles to photons today, since the number of black-body photons in a given co-moving volume today is approximately equal to the total number of particles in the gas at temperature T_R , due to the subsequent annihilation of these other particles.

If the shadow matter is to close the present-day Universe, its current energy density must be $\approx 10^4$ times larger than the current black-body radiation density. Let us assume that the N_s shadow matter particles all have equivalent masses M , an assumption that yields the largest remnant mass density today. Then the closure of the Universe with shadow matter requires

$$M \approx 10^{-8} (N_s^{1/2} / N_s) [M_{\text{Pl}} / T_R]^3 \text{ GeV} \quad (4)$$

Significant production of shadow matter by reheating requires $M < T_R$, which yields the bound

$$\begin{aligned} T_R^4 &\geq 10^{-8} (N_s^{1/2} / N_s) M_{\text{Pl}}^3 \text{ GeV} \\ \rightarrow T_R &\geq 10^{12} (N_s^{1/2} / N_s)^{1/4} \text{ GeV} \end{aligned} \quad (5)$$

This is our first result.

In the case that the shadow matter candidates for dark matter are heavier than the gravitino (or the gravitino is heavier than the reheating temperatures determined here), then, unless N_s is absurdly large, if there is inflation in the sector containing normal matter, the reheating temperature is constrained to be larger than $0(10^{12} \text{ GeV})$. In this limiting case, the mass of shadow matter particles must also be $0(10^{12} \text{ GeV})$. If the shadow matter is lighter, then the reheating temperature must be increased from this lower bound. As we discuss below, such reheating temperatures require either a gravitino mass, $m_g > 10^5 \text{ GeV}$ or $m_g < 0.1 \text{ GeV}$.

If shadow matter candidates are lighter than the gravitino, these particles may be produced by real rather than virtual intermediate gravitinos. In this case, due to the larger production cross-sections for gravitinos, we may expect that the abundance of shadow matter can be much greater, for lower reheating temperatures, than for the heavy shadow matter described above. We now demonstrate that while this is the case, cosmological constraints on gravitino number densities imply that a closure density of shadow matter can occur by this mechanism only if $m_g > 10^5 \text{ GeV}$.

The gravitino abundance relative to total particle number immediately after reheating is given approximately¹²⁻¹⁴ by

$$n_g / n \approx \alpha (T_R / M_{\text{Pl}}) \quad (6)$$

The resulting gravitinos can, on their decay, generate too much entropy¹¹⁻¹³, cause too large an abundance of stable photinos¹²⁻¹⁴ or lead to unacceptable photodissociation of light elements after nucleosynthesis¹⁵⁻¹⁷. The most stringent constraint on the reheating temperature for gravitinos comes from bounds on the light element abundance from photofission of ^4He by the decay photons from gravitinos. This yields the constraint^{16,17}

$$T_R < 0(10^{-9} [100 \text{ GeV} / m_g] \text{ GeV}) \quad (7)$$

If the energy from gravitino decay does not redshift (that is, it all remains in stable particles of mass m_g), then the reheating temperature required for the energy initially in gravitinos to close the Universe today is¹⁵

$$T_R \approx 1.3 \times 10^{12} [100 \text{ GeV} / m_g] \text{ GeV} \quad (8)$$

Thus, even if gravitinos deposited all their energy into the shadow matter sector and none of this energy were redshifted, according to the constraints on gravitinos from nucleosynthesis, there would not be enough energy (by a factor of 10^3 – 10^4) in shadow matter to close the Universe. (In fact, gravitinos will decay into light shadow-matter pairs at the same rate as they decay into normal matter pairs ($\sim m_g^3 / M_{\text{Pl}}^2$) so that unless $N_s \gg N_s$ then of the order of half the gravitino energy will be deposited in the hidden sector. We note in this regard that if most of the energy from gravitino decay went into the hidden sector, the upper bound in equation (7) would be raised. However, to raise it to be consistent with equation (8) would require $N_s > 1,000 N_s$, which again is not likely.) The gravitino lifetime is approximately M_{Pl}^2 / m_g^3 , so the gravitinos will decay before nucleosynthesis begins at $t \approx 1 \text{ s}$ only if $m_g > 10^5 \text{ GeV}$. In this case the constraint in equation (7) no longer applies.

Thus, whether shadow matter candidates are heavier or lighter than gravitinos, generation of sufficient shadow matter to close the Universe requires the gravitino mass to be $> 10^5 \text{ GeV}$ or $< 0.1 \text{ GeV}$. Such masses are inconsistent with minimal low-energy supergravity models which have $m_g \approx M_W$.

The case in which inflation occurs in the hidden E_8 sector is also strongly constrained by the requirement that sufficient energy be transferred to the observed sector. If the hidden sector is also assumed to provide the dark matter today, there are even further constraints.

If there are no significant out of equilibrium processes after reheating (including gravitino decay), then inflation in the hidden sector seems incompatible with observation. For, if we assume roughly equal numbers of helicity states in each E_8 sector, we find by interchanging s and s' in equation (3) that the ratio of the energy density in the observed sector to that in the hidden sector after reheating is given approximately by

$$\rho_0 / \rho_h \approx (T_R / M_{\text{Pl}})^3 \quad (9)$$

where now, of course, T_R refers to the reheating temperature of matter in the hidden sector. If there is no out of equilibrium reheating in the observed sector afterwards, then even if the hidden sector particles are all massless, this ratio will not change until well after nucleosynthesis, when the observed sector

becomes matter dominated. The energy density in the hidden sector must not contribute more than the effect of one extra massless neutrino at nucleosynthesis, so

$$\rho_h/\rho_0 \leq 1/6 \quad (10)$$

This requires that $N_s \leq N_2/6$ and that the reheating temperature T_R is near M_{Pl} . For such a temperature the approximations used to derive equation (9) break down. In any case, such high reheating temperatures are ruled out by considerations of generation of gravitational waves by inflation, and their effect on the microwave background and timing measurements of the millisecond pulsar¹⁹.

The only way for the ratio ρ_0/ρ_h to increase is for the observed sector to become matter dominated during a period when the hidden sector is not. This can occur through the existence of remnant long-lived massive particles. Suppose first that the decay of gravitinos provides this energy transfer between the two sectors. This decay must then occur before nucleosynthesis begins, at a temperature of the order of 1 MeV. One can show that as long as the pressure and energy density obey the inequality $0 \leq p \leq \rho/3$, then $d(H^{-1})/dt$ is bounded between 3/2 and 2. Thus, although this scenario involves strongly out-of-equilibrium decay, a temperature $T \approx 1$ MeV occurs on the order of the usual time of ≈ 1 s. Then, using the known decay rate for gravitinos, we again require $m_g > 10^5$ GeV. Furthermore, equation (10) implies that gravitino decays must release at least $\sim 85\%$ of their energy density into the observed sector. If the hidden sector is to provide the dark matter today, its relative energy density must be enhanced later as a result of its being matter dominated while the observed sector is radiation dominated.

The ratio of gravitinos to hidden sector particles immediately after reheating is given roughly by equation (6). To determine the minimum possible value of T_R , consider the borderline case in which $\rho_g \approx \rho_h$ at the decay time τ , in which case

$$n_g m_g \approx N_s T_h^4 \approx M_{Pl}^2/\tau^2 \approx m_g^6 M_{Pl}^2 \quad (11)$$

Using the fact that $n_h \approx N_s T_h^3$, one can compare the ratio obtained from equation (11) with that given in equation (6). This gives the bound

$$T_R > [m_g M_{Pl}]^{1/2} (N_s^{1/4} \alpha)^{-1} \quad (12)$$

Setting m_g equal to its minimum value and assuming $N_s^{1/4} \alpha \approx 10^{-2}$, one obtains $T_R > 10^{14}$ GeV.

Note that this reheating temperature could be reduced if gravitinos are produced by the decay of particles in the hidden sector, but could also be increased if the hidden sector has a matter-dominated era. Our discussion merely confirms that minimally acceptable scenarios for inflation in the hidden sector can be constructed. Of course, they would have severe problems with baryosynthesis, which we have not considered.

If the mass of the gravitino is of the order of 100 GeV, which is attractive from a model-building viewpoint, it does not decay fast enough to be responsible for the energy transfer into the observed sector. In this case, one must require that this is accomplished by other processes, and that gravitinos are themselves not produced at unacceptable levels.

In order to see whether this kind of a scenario is conceivable, we will assume that the energy transfer occurs with optimum efficiency. We therefore assume that the observed sector is dominated by N_M helicity states of 'M-particles' of mass $M \approx T_R$, with a lifetime τ chosen so that they decay at $T \approx 1$ MeV. The ratio r of these particles to hidden sector particles immediately after reheating is given analogously to equation (3) by

$$\Gamma \approx (N_M/N_s^{1/2}) T_R^3/M_{Pl}^3 \quad (13)$$

The ratio of the energy densities of the two sectors at this time is also given by this equation. If the hidden sector remains a radiation-dominated gas until after nucleosynthesis, then the ratio of the energy densities at $T \approx T_N \equiv 1$ MeV is increased

from r by a factor (T_R/T_N) . Requiring the ratio then to be of the order of unity, one finds

$$T_R > N_s^{1/8} N_M^{-1/4} (M_{Pl}^3 T_N)^{1/4} \approx 3 \times 10^{13} \text{ GeV} \quad (14)$$

This reheating temperature violates the photofission bound of equation (7) by five orders of magnitude, but this bound is irrelevant because in this scenario the particle number in the observed sector is increased by the factor $(T_R/T_N) \approx 3 \times 10^{16}$ between reheating and nucleosynthesis. This dilutes the gravitinos produced during reheating to a negligible density.

To be safe, however, we must still examine the production of gravitinos by the decay products of the M-particles. The number density of M-particles is given by

$$n_M(t) = N_0 (a(t_R)/a(t))^3 \exp\{-(t-t_R)/\tau\} \quad (15)$$

where $a(t)$ is the cosmological scale factor and t_R is the time of reheating. The initial number density N_0 of M-particles is given by $r N_s T_R^3$, and by using equation (13) and the minimal value of T_R from equation (14), one finds

$$N_0 \approx N_M^{-1/2} N_s^{5/4} (M_{Pl} T_N)^{3/2} \quad (16)$$

Assuming the energy released by the decay of the M-particles is rapidly thermalized, the energy density of observed sector radiation obeys

$$d(a^4 \rho_{\text{rad}})/dt = (a^4 n_M M)/\tau \quad (17)$$

The interesting time period will be $t \approx t_R \ll \tau$, so that we can take $\exp[-(t-t_R)/\tau] = 1$ and $a(t) \sim t^{1/2}$. With the initial condition $\rho_{\text{rad}}(t_R) = 0$, equation (17) can be solved to give

$$\rho_{\text{rad}} = 2 N_0 M t_R^{3/2} (t^{3/2} - t_R^{3/2}) \{3\tau t^2\}^{-1} \quad (18)$$

This expression is maximized at $t = 2^{2/3} t_R$, and has a maximum of the order of $N_0 M t_R/\tau \approx N_s^{5/8} N_M^{-1/4} (M_{Pl}^3 T_N^{1/4})^{1/4}$, corresponding to a temperature of about 10 GeV. The production of 100-GeV gravitinos is then completely negligible. Baryosynthesis, on the other hand, is clearly a severe problem.

We thus conclude that production of a non-negligible density of shadow matter in our Universe in the context of inflationary cosmology requires: (1) For inflation in the observed sector: either a reheating temperature in excess of 10^{12} GeV (implying $m_g > 10^5$ GeV or $m_g < 0.1$ GeV), combined with shadow matter particle masses given by the relation $M \approx 10^{12} \times \{(10^{12} \text{ GeV}/T_R)^3\}$ GeV, or $m_g > 10^5$ GeV and shadow matter masses less than m_g . (2) For inflation in the hidden sector: a reheating temperature in excess of $0(10^{13})$ GeV, combined with the presence of either new heavy particles in the observed sector which decay strongly out of equilibrium, or a gravitino with mass exceeding 10^5 GeV which deposits $>85\%$ of its decay energy in the observed sector.

Only the latter alternative, which has severe problems with baryosynthesis, allows a gravitino mass consistent with the simplest $N = 1$ low-energy supergravity models which, among other possibilities, could arise from a fundamental superstring theory.

L.M.K. is supported in part by the Harvard Society of Fellows, NSF grant PHY82-15249, at Harvard, and DOE grant 627 R 40073 at Yale. D.N.S. is an NSF predoctoral fellow. A.H.G. is supported DOE grant DE-AC02-76 ER03069, NASA grant NAGW-553, and the A. P. Sloan Foundation.

Received 27 September; accepted 30 December 1985.

- Green, M. B. & Schwarz, J. H. *Nucl. Phys.* **B181**, 502-530 (1981); **B198**, 441-461 (1982); *Phys. Lett.* **109B**, 444-448 (1982).
- Green, M. B. *Nature* **314**, 409-414 (1985).
- Gross, D. J., Harvey, J. A., Martinec, E. & Rohm, R. *Phys. Rev. Lett.* **54**, 502-505 (1985).
- Alvarez-Gaumé, L., Polchinski, J. & Wise, M. *Nucl. Phys.* **B221**, 495-523 (1983).
- Kolb, E., Seckel, D. & Turner, M. *Nature* **314**, 415-419 (1985).
- Guth, A. H. *Phys. Rev.* **D23**, 347-356 (1981).
- Linde, A. D. *Phys. Lett.* **108B**, 389-393 (1982).
- Albrecht, A. & Steinhardt, P. J. *Phys. Rev.* **Let.** **48**, 1220-1223 (1982).
- Abbott, L., Farhi, E. & Wise, M. *Phys. Lett.* **117B**, 29-40 (1982).
- Dolgov, A. D. & Linde, A. D. *Phys. Lett.* **116B**, 329 (1982).

11. Weinberg, S., *Phys. Rev. Lett.* **48**, 1303-1306 (1982).
12. Krauss, L. M. *Nucl. Phys. B* **227**, 556-569 (1983).
13. Weinberg, S. (unpublished).
14. Ellis, J., Hagelin, J., Nanopoulos, D. V., Olive, K. A. & Srednicki, M. *Nucl. Phys. B* **238**, 453-476 (1984).
15. Ellis, J., Kim, J. E., Nanopoulos, D. V. *Phys. Lett.* **145B**, 181-186 (1984).
16. Ellis, J., Nanopoulos, D. V., Sarkar, S., CERN preprint CERN-Th 4057/84.
17. Juskiewicz, R., Silk, J. & Stebbins, A. *Phys. Lett.* (in the press).
18. Ellis, J., Linde, A. D. & Nanopoulos, D. V. *Phys. Lett.* **118B**, 59-64 (1982).
19. Krauss, L. M. *Nature* **313**, 32-33 (1984).

Large-scale homogeneity of the Universe measured by the microwave background

N. Mandolesi*, P. Calzolari*, S. Cortiglioni*,
F. Delpino†, G. Sironi‡, P. Inzani‡, G. DeAmici‡¶,
J.-E. Solheim§, L. Berger§, R. B. Partridge||,
P. L. Martenis||, C. H. Sangreel|| & R. C. Harvey||

* TESRE/CNR, via de Castagnoli 1, 40126 Bologna, Italy

† Osservatorio Astronomico Universitario, 40126 Bologna, Italy

‡ Department of Physics, University of Milan and IFCTR/CNR, 20133 Milano, Italy

§ Institute for Mathematical and Physical Sciences, Tromsø N-9001, Norway

|| Haverford College, Haverford, Pennsylvania 19041, USA

The isotropy of the cosmic-microwave background is a measure of the homogeneity of the Universe at the epoch of last scattering, conventionally taken to be the epoch of recombination of the primeval plasma at a redshift $z_s \sim 1,000$. Here we report measurements of the isotropy at a wavelength of 3 cm. At the 95% confidence level, our results restrict departures from isotropy to $< 5 \cdot 7 \times 10^{-4}$ on angular scales of $\sim 2\text{--}10^\circ$. The angular scale of our observations is of particular cosmological significance because regions of size $\sim 1^\circ$ and above^{1,2} were not causally connected at the epoch of last scattering, at least in conventional big-bang models. Thus in these models, isotropy on scales $\geq 1^\circ$ can be stipulated only as an initial condition. In the new inflationary models³, on the other hand, a causal process to produce large-scale isotropy is included.

Searches for the quadrupole moment of the angular distribution of the background radiation^{4,5} have provided stringent limits on anisotropy, such as the limit by Fixsen *et al.*⁵ of $\Delta T/T \leq 3 \times 10^{-5}$ on angular scales $> 10^\circ$. The distribution of galaxies⁶, however, suggests that the amplitude of inhomogeneities in the Universe decreases with increasing scale. Hence, we believe it is necessary to check the isotropy of the background radiation at angular scales θ just above, but closer to, 1° . At present, the most closely relevant measurements are those reported at $\theta \sim 6^\circ$ by Melchiorri and co-workers⁷, who used a balloon-borne bolometric detector at an effective wavelength of ~ 0.7 mm. They detected anisotropy at the level $\Delta T/T = (3.7 \pm 0.7) \times 10^{-5}$, and believe that most of the effect they see may be ascribed to signals other than those arising from grey-body emission by dust clouds in our Galaxy. Recently, an upper limit of $\Delta T/T \leq 4 \times 10^{-4}$ on scales $\geq 5^\circ$ has been established by Strukov and Skulachev⁸ using a satellite-borne radiometer operating at $\lambda = 8$ mm. Other upper limits on scales $1\text{--}10^\circ$ available in the literature are either less sensitive⁹ or subject to question¹⁰⁻¹².

Before these results were available, we planned and began¹³ a search for anisotropies in the cosmic-microwave background on angular scales in the $2\text{--}10^\circ$ range at $\lambda = 3$ cm. We believe that observations at centimetre wavelengths are a valuable supplement to measurements at sub-millimetre wavelengths⁷ because the sources of possible systematic error are different

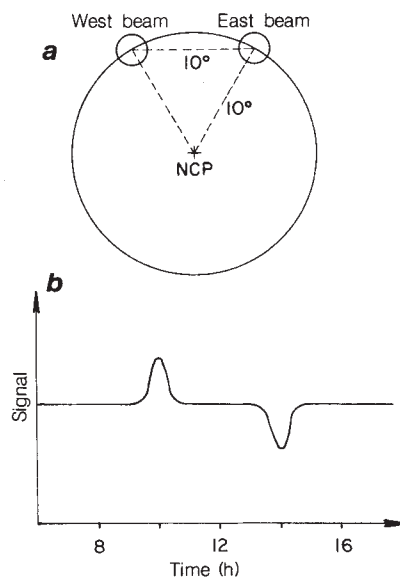


Fig. 1 a, Pattern of the beams on the celestial sphere; the large circle has declination $+80^\circ$ NCP; North Celestial Pole. b, Signature of a hotspot at 12 h passing through our beam pattern.

and because some models¹⁴ predict wavelength-dependent fluctuations.

Our radiometers were Dicke-switched, low-noise (system noise $\leq 1,000$ K), broad-band ($\Delta\nu = 0.5$ or 1.0 GHz), super-heterodyne receivers attached to small parabolic reflectors which gave us beams of $2\text{--}3^\circ$ full width at half maximum (FWHM). The two beams were directed at points 10° apart above the North Celestial Pole (NCP, see Fig. 1). As the Earth rotated, the beams traced a circle on the sky at a declination of 80° each sidereal day. Smaller horn antennas directed towards the NCP provided a useful reference and allowed us to subtract atmospheric emission to first order.

Solar heating and other diurnal effects, which could mimic anisotropy, were controlled by two different means. The first was to sum observations made over the course of an entire year, so that solar effects would average out in sidereal time. The second was to work where solar heating and daily temperature variations were absent or minimal.

The first approach was used with a radiometer located at Medicina, 30 km east of Bologna, Italy, which consisted of two identical, 1-m diameter, parabolic reflectors mounted equatorially to produce two identical beams of FWHM 2.3° . The radiometer input was mechanically switched between the two every 4 min of sidereal time. The receiver was calibrated every 4 days with a gas discharge tube seen through a directional coupler.

To keep the temperature of the system stable to 1°C , and to prevent condensation on the parabolas, the entire instrument was housed in a styrofoam dome with 20-cm thick walls, whose measured loss was $\leq 5 \times 10^{-4}$.

As expected, solar heating and atmospheric effects introduced both drift and solar-time variations into the recorded signals. Most of the variation was subtracted out, or was smoothed away, when we averaged data obtained during the 24 months of our runs. Nevertheless, equipment problems and poor winter weather reduced the data we received during the first observing year (1980) by 50%. We therefore used a second instrument in the Norwegian Arctic, where conditions are particularly suitable for winter observations. The site selected was a 750-m summit near Tromsø. The high latitude (70°N) ensured that our measurements were made at small zenith angles, reducing emission from the Earth's atmosphere. Most observations were made before the annual reappearance of the Sun over the local

¶ Present address: IRA/CNR, Bologna, Italy.

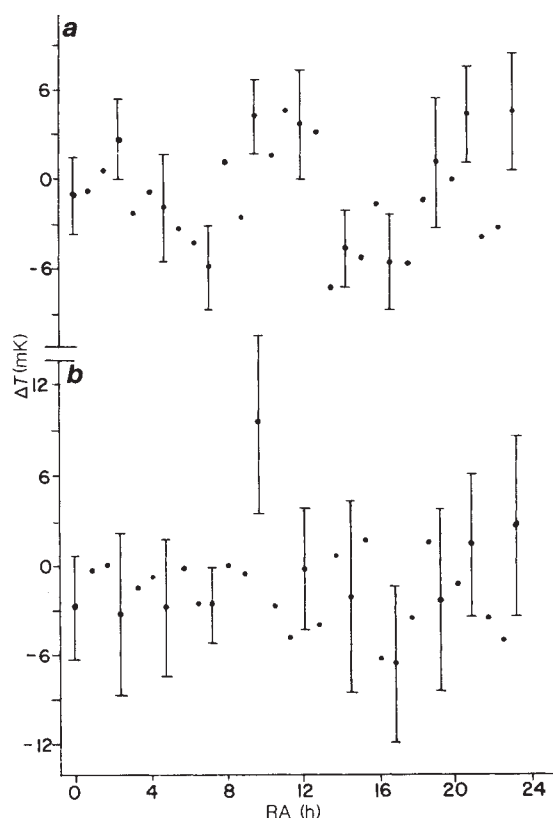


Fig. 2 Data from the two sites: *a*, Norway; *b*, Italy, averaged in 30 bins around the declination 80° circle. Typical error bars shown are $\pm 1\sigma$ of the mean.

Table 1 Measured brightness temperature differences as a function of position along the $\delta = 80^\circ$ circle

Bin no.	RA	Mean temperature (mK)	σ (mK)
1	0 h	-1.63	2.14
2	0 h 48 min	-0.81	1.39
3		0.64	1.63
4		1.63	2.39
5		-2.17	2.59
6	5 h	-0.76	2.72
7		-2.20	2.82
8		-2.58	2.59
9		-3.78	2.22
10		-4.04	1.86
11	8 h	0.84	1.81
12		-2.32	1.73
13		5.06	2.26
14		0.39	2.34
15		2.69	2.57
16	12 h	1.98	2.65
17		0.39	2.53
18		-4.06	3.10
19		-4.20	2.42
20		-2.75	3.08
21	10 h	-2.81	3.57
22		-5.79	2.71
23		-5.10	3.01
24		-0.46	3.41
25		-0.02	3.51
26	20 h	-0.26	3.07
27		3.49	2.68
28		-3.72	3.30
29		-3.23	1.27
30	23 h 12 min	4.02	3.23

The weighted averages of the combined Italian and Norwegian measurements are shown. σ , standard deviation of the mean.

horizon (~ 10 March). Diurnal variations in the ambient temperatures were small (often $< 3^\circ\text{C}$), and the radiometer was housed in a thermostatically controlled enclosure.

Two identical feed horns symmetrically placed in the focal plane of a single 1.2-m diameter parabolic reflector produced two beams 10° apart (as in Fig. 1). The radiometer input was switched from one beam to the other every 14 s by a low-loss mechanical switch near the focus. Illumination of the reflector was tapered to produce beams with low sidelobes and approximately gaussian shape of FWHM 2.8° . A ferrite Dicke switch alternated the radiometer input at $\sim 1,000$ Hz between one of the two feed horns and a small reference horn.

The instrument was calibrated using a sheet of ambient temperature Eccosorb. A 13% correction was included to account for the non-linearity of the second detector in the receiver. More frequent secondary calibrations were performed by changing the insertion loss between the reference horn and the Dicke switch. The sensitivity, efficiency and antenna pattern of both instruments were checked by making drift scans of the Sun and Moon.

Our data consisted of differences in antenna temperature between a reading in the east beam and a reading in the west beam. These differences were grouped into 30 bins in solar time, each 48 min wide. The number of samples in each bin was determined by the switching time; ≤ 103 for the Norwegian work or 12 for the Italian. Samples exceeding 5 times the standard deviation of the values in a given bin were dropped.

Next, a line was fit to the mean values of the solar time bins in each run, and then subtracted to eliminate long-term drift and some of the diurnal temperature variation. This procedure reduced the number of degrees of freedom in our data, and made us insensitive to very large angular scale fluctuations in

the sky temperature. There are 2 degrees of freedom lost in the Norwegian data; because the individual Italian runs were < 24 h in length, more degrees of freedom (up to 6) were lost in these data.

After the linear fit, the residuals were rebinned in sidereal time, and averaged with data from all other runs. The weight used for averaging the Norwegian results was the inverse of the variance of the data in that bin; for the Italian runs, the observing time. The results of this weighted averaging are shown in Fig. 2 (see also Table 1).

Do these results provide evidence for real anisotropy in the microwave background? First, note that any region of enhanced temperature along the $\delta = 80^\circ$ circle would produce a positive signal when passing through the east beam, then a negative signal of equal amplitude 4 h (or 5 bins) later when passing through the west beam. Second, real anisotropy would appear in both the Italian and the Norwegian data. An inspection of Fig. 2 shows no such evidence above a level of ~ 4 mK.

Tighter constraints may be set by comparing the variance of the observed results with the variance introduced by instrumental and atmospheric noise alone. To achieve this, we use the statistical measures advocated by Lasenby and Davies⁹. The first is an unbiased⁹ estimator of the sky variance $\hat{\sigma}^2$. Taking account of losses of degrees of freedom from the polynomial fit and our beam pattern, we find $\hat{\sigma}^2 = 1.90$, -0.85 and $1.54 (\text{mK})^2$, for the Norwegian, Italian and combined data, respectively. For measurements involving the difference between two independent samples of the sky, the r.m.s. value of the sky fluctuation σ_{sky} is related to $\hat{\sigma}^2$ by $\sigma_{\text{sky}} = \sqrt{\hat{\sigma}^2/2}$. If we now make the conservative assumption that all the sky fluctuation is produced by anisotropies in the cosmic microwave background,

the fractional fluctuation is $\Delta T/T = \sigma_{\text{sky}}/2.7$ as shown in Table 2.

Lasenby and Davies⁹ also point out that overestimation of the errors of the data points will bias the estimates of $\hat{\sigma}$ downward. Thus, they suggest adopting the larger of two other statistical measures c and s_{ii}^2 as a more conservative upper limit on $\hat{\sigma}^2$. The first involves a test of the hypothesis that $\sigma_{\text{sky}} > 0$ against $\sigma_{\text{sky}} = 0$; if $\hat{\sigma}^2 > c$, we infer $\sigma_{\text{sky}} > 0$. The entries in Table 2, row 2 are for a test at the 95% significance level. The second quantity is a 95% upper confidence level on $\hat{\sigma}$ (see Table 2, row 3) and is related to the statistic used by Boynton and Partridge^{15,16} which produces similar limits on σ_{sky} for our data.

In the case of the Norwegian data, the value of $\Delta T/T$ calculated from $\hat{\sigma}^2$ is slightly larger than that found from the measure c , suggesting the possibility that real fluctuations are present. As this result is not confirmed in the Italian data, and is barely significant, we prefer to treat our data as upper limits on $\Delta T/T$ on angular scales of ~ 2 – 3° . Using the combined data set, we have $\Delta T/T \leq 5.6 \times 10^{-4}$ at 95% confidence limits.

Table 2 Upper limits at 95% confidence on $\Delta T/T = \sigma_{\text{sky}}/2.7$ determined from the observations using the statistical measures of ref. 9

Statistical measure used	$\Delta T/T (\times 10^{-4})$			
	Norwegian data	Italian data	Combined data Unsmoothed $n=1$ ($\theta \approx 2^\circ$ – 3°)	Smoothed $n=2$ ($\theta \approx 5^\circ$)
$\hat{\sigma}^2$	3.6	2.4	3.2	3.5
c	3.5	3.0	3.1	3.0
s_{ii}^2	6.2	1.6	5.6	6.8

We may also smooth our observations by taking a running average of pairs of points (Table 2). Finally, note that for comparison with theoretical predictions of the angular correlation function of temperature fluctuations, the appropriate angular scale is 10° , our beam-switching angle.

Although not as sensitive as some results reported from shorter wavelength or larger-scale measurements^{4,5,7}, our results show either that some causal process acted to smooth the Universe on large scales before the epoch of last scattering (as in the inflationary models), or that the matter distribution in the Universe was very homogeneous *ab initio*. For instance, for large-scale adiabatic^{1,2} perturbations, we expect $\Delta T/T \sim 1/3 \Delta \rho/\rho$; hence our results imply $\Delta \rho/\rho \leq 2 \times 10^{-3}$ for density perturbations at $z \sim 1,000$.

We thank the Consiglio Nazionale delle Ricerche, NSF, and NATO research grant 1871 for partial support of this work. R.B.P. was supported by a Fulbright fellowship while in Tromsø. We also thank Troms Kraftforsyning and Høyer-Ellefsen for assistance in Norway, and especially Ole Kristian Hegge for hospitality. K. E. Egge and R. B. Kalleberg assisted in the observations.

Received 22 July; accepted 30 December 1985.

1. Peebles, P. J. E. *Physical Cosmology*, Ch. 5 (Princeton University Press, 1971).
2. Weinberg, S. *Gravitation and Cosmology* Ch. 15 (Wiley, New York, 1972).
3. W. & Siklos, S. T. C. (eds) *The Very Early Universe* (eds Gibbons, G. W., Hawking, S. W. & Siklos, S. T. C.) (Cambridge University Press, 1983).
4. Lubin, P. M., Epstein, G. L. & Smoot, G. F. *Phys. Rev. Lett.* **50**, 616–619 (1983).
5. Fixsen, D. J., Cheng, E. S. & Wilkinson, D. T. *Phys. Rev. Lett.* **50**, 620–622 (1983).
6. Peebles, P. J. E. *The Large Scale Structure of the Universe* (Princeton University Press, 1980).
7. Melchiorri, F., Melchiorri, B., Ceccarelli, C. & Pietranera, L. *Astrophys. J. Lett.* **250**, L1–L4 (1981).
8. Strukov, I. A. & Skulachev, D. P. *Soviet Astr. Lett.* **10**, 1–4 (1984).
9. Lasenby, A. N. & Davies, R. D. *Mon. Not. R. Astr. Soc.* **203**, 1137–1169 (1983).
10. Parijskij, Yu. N., Petrov, Z. F. & Cherkov, L. N. *Soviet Astr. Lett.* **3**, 262–265 (1977).
11. Berlin, A. B. *et al.* *IAU Symp.* No. 104, 1983.
12. Lasenby, A. N. thesis, Manchester Univ. (1981).
13. Calzolari, P. *et al.* *17th Int. Cosmic Ray Conf.* **9**, 270–271 (1981).
14. Hogan, C. J. *Mon. Not. R. Astr. Soc.* **192**, 891–903 (1980).
15. Boynton, P. E. & Partridge, R. B. *Astrophys. J.* **181**, 243–254 (1973).
16. Partridge, R. B. *Astrophys. J.* **235**, 681–687 (1980).

Is there an unusual solar core?

Ken G. Libbrecht

Solar Astronomy 264-33, California Institute of Technology, Pasadena, California 91125, USA

Various attempts have been made in the past two decades to determine the properties of the core of the Sun, by measurements of solar neutrino radiation, the distortion of the solar gravitational potential and frequency splittings of solar oscillation modes. In the latter category, the two best measurements have been made by Duvall and Harvey¹ and by Brown²; both showed splitting roughly independent of spherical harmonic degree l , and both had a peculiar peak in the measured splitting at $l = 11$. We present here new results, based on the analysis of 6,656 individual oscillation modes for $5 \leq l \leq 20$. These data yield a splitting spectrum which is consistent with previous measurements, but without the unusual peak at $l = 11$, thus suggesting that a simple standard model for the solar core is essentially correct.

Several recent measurements have been interpreted as evidence for an unusual solar core, where unusual here means either rotating rapidly relative to the Sun's surface or containing very strong magnetic fields.

The first accurate measurement of the shape of the Sun in 1966 by Dicke and Goldenberg³ yielded an oblateness⁴ of $\Delta r \equiv r_{\text{equator}} - r_{\text{pole}} = 41.9 \pm 3.3$ arc ms (or a solar gravitational quadrupole moment $J_2 \times 10^6 = 23.7 \pm 2.3$), much larger than the $\Delta r = 7.8$ arc ms caused by surface rotation alone. Further analysis of the same data revealed a 12.4-day rotating distortion which was interpreted as the signature of a rotating magnetic core in the Sun⁵. However, the measurement was repeated^{6,7} in 1983 and a much smaller oblateness of $\Delta r = 15.4 \pm 4.1$ arc ms was found (after taking into account any possible latitude-dependent brightness signal⁶), giving $J_2 \times 10^6 = 5.3 \pm 2.8$, and no significant 12-day oscillations were observed. Analysis of new 1984 data from the distortion telescope has yielded similar results, with a smaller oblateness measured for that year²³.

Claverie *et al.*⁸ detected a 12.6-day oscillation in the line-of-sight velocity of the integrated solar surface, which was interpreted as evidence of a rotating magnetic core. However, the observations were later found to be explained by sunspots travelling across the solar disk^{9–11}.

A triplet structure with a spacing of $0.75 \mu\text{Hz}$ was found in $l = 1$ p-modes by the Birmingham group¹², which was initially taken to be rotational¹² or magnetic^{13–15} splitting of the different m -states of the $l = 1$ modes. However, it now seems that the fine structure in the power spectrum could be a manifestation of the short lifetimes of the p-modes, and not a true splitting^{16–18}. Analysis of 290 days of solar irradiance data collected by the Solar Maximum Mission spacecraft^{16,17} indicated that linewidths for $l = 0, 1, 2$ modes are typically $1.5 \mu\text{Hz}$, and a 2.5σ upper limit of $1.0 \mu\text{Hz}$ was found for the splitting between adjacent m -values of $l = 1$ modes from line-broadening measurements. This implies that $J_2 \times 10^6 < 1.0$ from rotation.

Hill¹⁹ recently measured p-mode splittings of $1.8 \pm 0.1 \mu\text{Hz}$ for $1 \leq l \leq 6$ modes, which is greater than other measurements and is especially inconsistent with the results of Duvall and Harvey and the present work. One possible explanation for the discrepancy may lie in Hill's data set, which measured only three solar diameters on the Sun at each time interval and is therefore sensitive to all modes with $l < 50$ (ref. 19). That proper mode identification has been made for the $\sim 10^4$ modes with frequencies from 2 to 4 mHz in a single power spectrum has yet to be established.

Duvall and Harvey¹ measured the splitting of $m = \pm l$ p-modes for $1 \leq l \leq 100$, finding a nearly constant splitting of $0.46 \pm 0.02 \mu\text{Hz}$. The splitting decreased slightly with decreasing l , with

a peak at $l=11$, then turned over at $l \approx 4$ to further increase with decreasing l , but with quite large errors for smaller l . Inversion of this splitting spectrum²⁰ gave a solar internal rotation that is approximately constant with radius down to $r=0.3 R_{\odot}$, giving $J_2 \times 10^6 = 0.17 \pm 0.04$. Note that this value of J_2 is calculated from a derived rotation versus depth curve, unlike the direct measurement of J_2 from the solar shape measurement³. It is still possible that a large magnetic field may distort the mass distribution of the solar interior, giving a larger J_2 , although how such a field would be stabilized is uncertain.

Brown² recently made another measurement of the splitting for modes of all m -states with $8 \leq l \leq 50$, finding splittings slightly larger than the Duvall and Harvey numbers for $10 < l < 30$, and also nearly exactly the same peak at $l=11$. Neither peak by itself could be claimed significant, but the appearance of a single nearly identical peak in two different data sets was suggestive of an unusual solar core².

The splitting measurements by Duvall and Harvey used measured individual mode frequencies to determine splittings, but their observational technique of optically averaging the solar image allowed some leakage of modes with $m < l$ into their data, which affected the derived splittings. Problems of this sort were considered in their analysis¹. Brown used full two-dimensional spatial filtering to isolate modes, which virtually eliminates this problem from the outset. However, the many modes considered in that analysis necessitated a simple cross-correlation approach to measure the splittings². I have found that such an approach is a poor substitute for the use of isolated individual mode frequencies. To address better the question of the splitting of low- l oscillation modes, I have investigated the 6,656 individual modes for all m -values with $5 \leq l \leq 20$ and having radial order $9 \leq n \leq 24$.

In 1984-85, a new telescope was built at Big Bear Solar Observatory, and is now dedicated to helioseismology observations. Using a Zeiss 0.25-Å bandpass birefringent filter chopped ± 0.125 Å by an electro-optical crystal, solar images in the two wings of the 6,439-Å Ca line were alternately formed on an RCA Newvicon TV camera. A COSMOS image processor digitized and added/subtracted the blue wing/red wing images at the rate of 7.5 pairs per s, and added 375 pairs to produce one 192×240 -pixel Doppler image of the Sun each minute. For our present analysis, data from 12 contiguous days of observations, 22 July 1985 to 2 August 1985, were scaled, rotated and collapsed to form 8,150 32×32 -pixel images. After subtracting a running mean image to remove slow trends, the residuals were fitted to individual spherical harmonics with $5 \leq l \leq 20$ and the resulting 416 time series of fit coefficients were padded with zeros and Fourier-transformed.

Individual modes were easily identified using existing tables²¹, and the mode frequencies were fit to a polynomial in l and n as a first step in the analysis. Using the fit polynomial as a guide, gaussian curves were fit to the individual modes and the l - n polynomial fit was redetermined, leaving residuals of ± 2 μHz. Mode frequencies were further refined by then refitting the gaussians and by using a CLEAN algorithm¹ on 10-μHz-wide segments of the power spectra centred on the fit frequencies. The mode frequencies determined from both these techniques were combined, and each was assigned a weight using its fit gaussian width, the second moment of its CLEANed spectrum, and the ratio of power in centred 10- and 50-μHz sections of its CLEANed spectrum. The mode frequencies for each l were then fit by weighted least-squares to

$$\nu_{nlm} \rightarrow R_{nl} + A_l m + B_l m^2 - C_l m^3$$

Figure 1 shows the results of the fit, which are in good agreement with Brown's² data. The larger noise at small l is caused by the decreasing total number of modes available for the fit; all three plots are consistent with there being no l -dependence in the data. For comparison with both Brown and Duvall-Harvey the data have been redisplayed in Fig. 2 as the rotational splitting

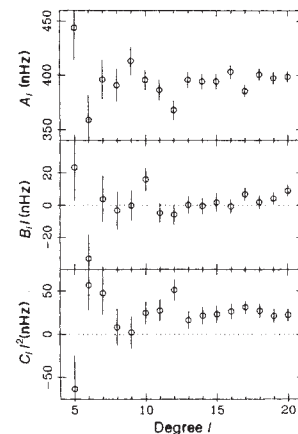


Fig. 1 Linear, quadratic and cubic fit coefficients described in the text. Factors of l and l^2 have been included to remove most of the l -dependence. The fits were done from a synodic reference frame.

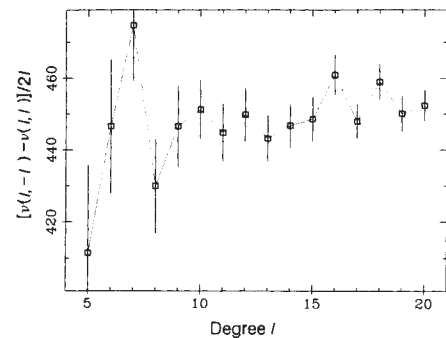


Fig. 2 Rotational splitting of oscillation modes plotted against degree, averaged over radial order n for each l . $A_l + C_l l^2 + 31$ is plotted from the fit, where the 31 nHz is added to shift to a sidereal reference frame.

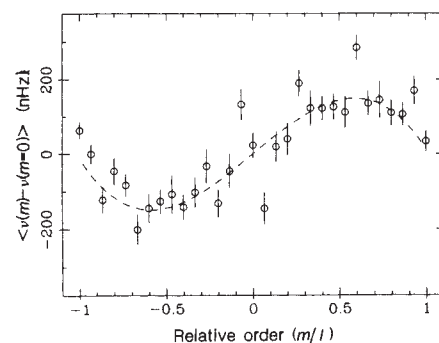


Fig. 3 Dependence of oscillation frequency on order m , where all modes with $10 \leq l \leq 20$ and $9 \leq n \leq 24$ have been combined, and a linear dependence of 420 m nHz (synodic) has been subtracted. Dashed line, best-fit cubic $\nu(m) = 381(m/l) - 369(m/l)^3$.

plotted against l . Here, it is clearly seen that the 30-nHz peak at $l=11$ shown by both previous splitting measurements is not visible in the Big Bear data set. Assuming that no real peak exists in the splitting spectrum, one can conclude from these results and the above-mentioned results of other solar researchers that there is little or no evidence at present for a rapidly rotating and/or magnetic solar core.

The Big Bear data also show the solar differential rotation (Fig. 3). This plot and the cubic fit coefficients C_l in Fig. 1 are in good agreement with Brown's results.

The measured splitting in Fig. 2 is roughly constant at 450 nHz, which is ~ 10 nHz larger than the same region measured

by Duvall and Harvey. This may be an artefact of the different methods of data analysis, or some small mode leakage in the Duall-Harvey data¹; or it may be a physically significant shift. Howard²² used 62 yr of Mt Wilson sunspot data to measure averaged systematic changes of ~ 5 nHz in the Sun's surface rotation rate at different times in the solar cycle. We may now be seeing a related shift in 3 yr of helioseismology data.

This work was supported by NSF grant no. NSF ATM-8320435.

Received 27 September; accepted 30 December 1985.

1. Duvall, T. L. Jr & Harvey, J. W. *Nature* **310**, 19–22 (1984).
2. Brown, T. M. *Nature* **317**, 591–594 (1985).
3. Dicke, R. H. & Goldenberg, H. M. *Astrophys. J. Suppl.* **27**, 131–182 (1974).
4. Dicke, R. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1309–1312 (1981).
5. Dicke, R. H. *Solar Phys.* **78**, 3–16 (1982).
6. Dicke, R. H., Kuhn, J. R. & Libbrecht, K. G. *Nature* **316**, 687–690 (1985).
7. Libbrecht, K. G. *thesis*, Princeton Univ. (1984).
8. Claverie, A. *et al.* *Nature* **299**, 704–707 (1982).
9. Durrant, C. J. & Schröter, E. H. *Nature* **301**, 589–591 (1983).
10. Andersen, B. N. & Maltby, P. *Nature* **302**, 808–810 (1983).
11. Edmunds, M. G. & Gough, D. O. *Nature* **302**, 810–812 (1983).
12. Claverie, A., Issak, G. R., McLeod, C. P. & van der Raay, H. B. *Nature* **293**, 443–445 (1981).
13. Isaak, G. R. *Nature* **296**, 130–131 (1982).
14. Dicke, R. H. *Nature* **300**, 693–697 (1982).
15. Gough, D. O. *Nature* **298**, 350–354 (1982).
16. Woodard, M. & Hudson, H. S. *Nature* **305**, 589–593 (1983).
17. Woodard, M. *thesis*, Univ. California San Diego (1984).
18. Grec, G., Fossat, E. & Pomerantz, M. *Solar Phys.* **82**, 55–66 (1983).
19. Hill, H. A. *Astrophys. J.* **290**, 765–781 (1985).
20. Duvall, T. L. Jr *et al.* *Nature* **310**, 22–25 (1984).
21. Harvey, J. W. & Duvall, T. L. Jr *Proc. Snowmass Conf. Solar Seismology from Space* (ed. Ulrich, R. K.) (Jet Propulsion Laboratory, Pasadena, 1984).
22. Howard, R. *An. Rev. Astr. Astrophys.* **22**, 131–155 (1984).
23. Dicke, R. H., Kuhn, J. R. & Libbrecht, K. G. *Astrophys. J.*, submitted.

Mode decoupling during retrorefraction as an explanation for bizarre radar echoes from icy moons

Von R. Eshleman

Center for Radar Astronomy, Stanford University, Stanford, California 94305, USA

Radar echoes from Europa, Ganymede and Callisto, large ice-covered moons of Jupiter, have characteristics that differ markedly from those of other natural targets^{1–3}. A decoupling of two characteristic modes of propagation appears to be the key to an understanding of the anomalous distribution of power among the polarized components of the echoes. Such decoupling is inherent in the double-reflection theory of Ostro and Pettengill⁴ and the sub-surface scattering model of Goldstein and Green². However, neither of these theories explains the astounding strengths of the echoes as well as the retrorefraction peak in the refraction scattering model of Hagfors *et al.*⁵, which in turn is less able to represent the power distribution. The two concepts of mode decoupling and retrorefraction are combined here to explain the strength, distribution and other features of the echoes. Causes of the decoupling are proposed and their morphological implications are introduced.

Important aspects of the problem are illustrated in Table 1. The first column gives the polarized and normalized radar cross-sections measured at wavelength 0.13 m for Europa, from the review by Ostro⁶. In each case a polarized wave, either linear (L) or circular (C), is transmitted from Earth and echo power is measured in the same (S) and orthogonal (O) senses of polarization. Note how the numbers for Europa compare with the next column of representative values for rocky planets and our Moon, the echoing properties of which are relatively well understood. Particularly striking are the Europa cross-sections which exceed those of a perfectly reflecting metal sphere, and the enormous returns in the 'unexpected' OL and SC polarizations where smooth reflecting surfaces would give zero values. These are the principal features to be explained by theory.

In their treatment of refraction scattering, Hagfors *et al.*⁵ envisage a large number of centres of refraction beneath or on the surface, away from which the refractive index decreases to merge with a smaller background value representing the bulk of the near-surface material of the moon. Radar ray paths are conducted around or bent toward these centres as in a lens, and those rays that are deflected by π radians make up the echoes that are received back on Earth. Hagfors *et al.*⁵ use a geometrical optics (GO) treatment for spherical lenses which predicts infinite retrorefraction with the echo power distributed among polarizations, as given in Table 1 (GO sphere). This model is successful in illustrating strong echoes and the large OL and SC components in particular. The authors invoke more standard scattering mechanisms to explain differences between theory and observation for the relative strengths of the returns. For example, they must ascribe all of the observed OC echo to such sources. However, anomalously large cross-sections are observed for all polarizations. If this model can produce only three large values, there remains the need for a second special source for the fourth. The quest here is for a single explanation of all four.

First consider the meaning of the GO infinities: they can be removed by the quasi-wave-optical (WO) method that has been used for the forward caustics in radio occultations by planetary atmospheres⁷ and gravitational lenses⁸. For retrorefraction this yields

$$\sigma/\pi a^2 = (16\pi a/\lambda q)P \quad (1)$$

where σ is the total radar cross-section; a the ray impact parameter which is the same as the radius of the coherent aperture and is assumed to be large relative to the wavelength λ ; q is a numeric defined at $\alpha = \pi$ by $|da/d\alpha| = a/\pi q$ where α is the angle of refraction; and $0 \leq P \leq 1$ is a polarization factor to be discussed. Equation (1) corresponds to antenna-like radiation from the coherent annular aperture shown in Fig. 1A having radius a and strip width $(\lambda a/\pi q)^{1/2}$. The (a/λ) factor in equation (1) shows promise for explaining strong echoes, if P is not too small.

But that is a problem. Figure 1A illustrates that $P = 0$ for all retrorefracting lenses that are perfect spheres; that is, a polarized incident wave emerges from the coherent annular aperture with components that exactly cancel each other. Thus, the infinities of GO become zeros of the quasi-WO solution. The theory appears to be as bizarre as the observations. However, $P = 4\sqrt{2}/\pi^2$ at $\alpha = \pi - \lambda/2a$ so there is a cone of strong refraction with a null at $\alpha = \pi$. If the lenses were imperfect, the null could be filled with energy from this adjacent conical lobe.

For an imperfect spherical or spheroidal lens, the aperture of Fig. 1A is no longer coherent. Ray paths emerging at substantially different azimuths are of randomly different lengths so the annulus breaks up into local patches or glints of coherence, as illustrated in Fig. 1B. If there were more than about 2π patches, the polarization would be essentially constant over their individual areas and the echo power from the individual glints would be summed. It follows that for a loss-less spheroidal lens

$$\sigma/\pi a^2 = T = (16F/q)(H/\lambda) \quad (2)$$

where H is the circumferential length of a representative coherent patch and F the fraction of azimuths covered by such patches. With these concepts, $F < 1$ and $\lambda < H < a$. Consideration of the polarization of the refracted rays, as in Fig. 1B, yields the 'WO spheroid' column of Table 1. This evaluates the GO infinities but still leaves the zero for the OC echo.

In finding echo polarizations in the above example, each incident ray was broken down into two modes, a transverse electric (TE) mode (electric vector normal to the plane of the refracted ray in the lens) and a transverse magnetic (TM) mode (electric vector in the plane of propagation). It was assumed that these modes were affected in the same way by the lens, except for the geometrical reversal of the electric field of the TM mode, before they were reassembled at the lens exit into

Table 1 Theoretical values of $\sigma/\pi a^2$ compared with observed values of $\sigma/\pi R^2$

	$\sigma/\pi R^2$ from observation			$\sigma/\pi a^2$ from theory of lens retrorefraction					Example ($4A=2.60$, $x=0.230$)
	Icy Europa	Rocky planet	Metal sphere	GO sphere	WO sphere	WO spheroid	WO spheroid (zero mode)	WO spheroid (changed modes)	
SL	1.8	0.09	1.0	$\infty/2$	0	$T/2$	$3T/8$	$A(3-x)$	1.80
OL	0.8	0.01	0	$\infty/2$	0	$T/2$	$T/8$	$A(1+x)$	0.80
SC	1.6	0.02	0	∞	0	T	$T/4$	$2A(1+x)$	1.60
OC	1.0	0.08	1.0	0	0	0	$T/4$	$2A(1-x)$	1.00
S+O	2.6	0.10	1.0	∞	0	T	$T/2$	$4A$	2.60

σ , Radar cross-section for backscatter; R , radius of planet, moon or metal sphere; a , lens radius at impact parameter yielding retrorefraction; S , echo energy polarized the same as that of the incident wave; 0, echo energy polarized orthogonal to that of the incident wave; L , incident linear polarization; C , incident circular polarization; ∞ , theoretical infinity suggesting a large value; GO, geometrical optics theory; WO, quasi-wave-optics theory; T , total normalized cross-section from equation (2); $4A = (T/2)d^2(1+e^2)$; $x = 2e \cos \theta / (1+e^2)$, where θ = electric phase-angle difference of characteristic modes; d , e , mode field-strength ratios discussed in text.

SL and OL or SC and OC components. But there are mechanisms that could differentially change the amplitudes, phases and paths of the modes during refraction, and this would redistribute the power. For example, if the TE mode were transmitted through the lens without loss while the TM mode were totally lost, Fig. 1B would be changed to Fig. 1C and there would be a corresponding change for incident circular polarization; this would result in the distribution given in the 'zero mode' column of Table 1 which shows, as desired, four strong components.

Further analysis of the same type has been used for a differential phase shift of θ radians between the two modes, and field strength ratios of d and e where $e \leq 1$ is the ratio of the weaker to the stronger field after refraction of TM and TE input fields of equal strength, and $d \leq 1$ is the ratio of the stronger field to what it would have been if there had been no loss during refraction. In general, there would then be independent TM and TE glints, as in Fig. 1D. Polarization results are given in the 'changed modes' column of Table 1, with total echo strength $4A = (T/2)d^2(1+e^2)$ where T is derived from equation (2), and with the mode decoupling factor $x = 2e \cos \theta / (1+e^2)$. The final column of Table 1 provides an example with $4A = 2.60$, representing Europa, and a selected value of $x = 0.23$ for optimum matching. The first and last columns therefore constitute a comparison of observation and theory.

Table 2 shows a comparison of theoretical values with observations of all three moons. The mean observed results⁶ are quoted to two decimal places, which is better than the expected measurement precision, and each of the theoretical comparisons involves normalization by the total (S+O) cross-section observed for that moon. The remaining two degrees of freedom for each moon then means that there are six degrees of freedom in the observational array being compared with a theoretical array having only one degree of freedom.

Table 2 compares whole moons with a single retrorefracting lens. To the extent that a good match is obtained, the implication is that essentially all of the observed echo power could be due to a large number of lenses having a distribution of x values averaging 0.23. Very little or no contribution would then be expected from other scattering mechanisms that partition the polarized power in a different manner. The enormous echo strengths could be modelled, for example, with lenses of 20 m radius covering 4% of the surface area of Europa, if $a = 10$ m, $H = 2$ m, $\lambda = 0.13$ m, $q = 1$, $F = 0.5$ and $d = e = 1$.

Neither single reflection nor refraction alone can explain the echoes. In a sense, the x parameter represents a continuum that lies between reflection and refraction, with $x = -1$ for the anti-coupled modes of pure reflection, $x = +1$ for the coupled modes of refraction, and $x = 0$ for the midpoint of complete mode decoupling. These conditions are represented in Table 1, by the metal sphere, WO spheroid, and zero-mode columns, respectively. Note that $x = 0$ means $e = 0$ or $\cos \theta = 0$, where in the latter case either $\theta = \pm \pi/2$ or θ is randomized over 2π so that

the average value of $\cos \theta$ is zero. The particular value of $x = 0.23$ used in Table 2 could be caused by a combination of deterministic amplitude and phase differences, or by slightly less than complete phase randomization.

The lenses considered here have refractive indexes $n(r)$ that generally decrease with radius r . I defer to Hagfors *et al.*⁵ for a discussion of how this might arise in the environment of the icy moons, but add some comments related to decoupling. Several undulations or discontinuities in $n(r)$, added to its general decrease with radius, could cause mode trapping over intervals of the ray paths, as in a dielectric waveguide or optical fibre. There would be an electrical phase difference between TM and TE modes at the guide output, and effective guides of different lengths could randomize the relative phase. Other models involving 'leaky' waveguide effects and localized dielectric structures of high loss have also been considered for decoupling. A promising model for future study is one in which mode trapping is responsible for both decoupling and backward focusing, even in the absence of a general refractive gradient.

Undulations or discontinuities in $n(r)$ might result from centres of first or last freezing, where the process of solidification causes a separation of material by concentration and by chemical

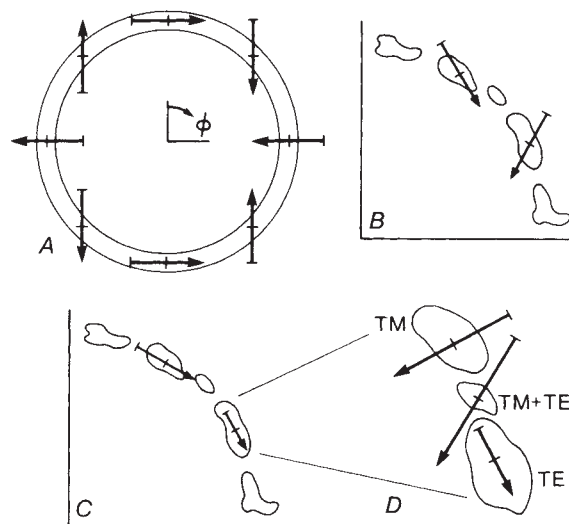


Fig. 1 Coherent apertures (outlined zones) and polarizations (arrows) of retrorefracted waves for horizontal incident polarization, as viewed from the position of the radar. **A**, Spherical annular aperture of radius a and strip width $(\lambda a / \pi q)^{1/2}$. **B**, Glints from one-quarter of a spheroidal lens with no mode differentiation. **C**, Glints from one-quarter of a spheroidal lens with zero TM mode. **D**, Glints near $\phi = \pi/3$ from a spheroidal lens with almost independent TM and TE components.

composition. The rate of freezing in a diurnal cycle might help to create undulations. Crystallization in the form of radial spikes, and surface flooding with subsequent sublimation of ice to concentrate entrained dust in thin layers, are other mechanisms related to the above models. Half-spheroids could also produce retrorefractions: here one might envisage mode-trapping structures produced during cratering, or an original liquid surface layer that left separate and roughly hemispherical pools after partial freezing. The final freeze might then create structures with $n(r)$ increasing toward the centre and with discontinuities and possible loss layers being superimposed.

There is a second-order effect not included in the above theory that is due to non-planar refraction, or refraction with a 'twist' of the modes. If this effect were taken into account, perfect agreement with the observations could be obtained for each moon by the proper choice of x and a characteristic mean-square value of random rotation. This refinement could be of significance for Callisto, which shows the poorest fit in Table 2.

Table 2 Observed and theoretical values of normalized radar cross-sections for three icy moons of Jupiter

	Observations			Theoretical values $x = 0.23$		
	Europa	Ganymede	Callisto	Europa	Ganymede	Callisto
SL	1.77	1.03	0.41	1.80	1.05	0.44
OL	0.83	0.49	0.23	0.80	0.47	0.20
SC	1.58	0.92	0.35	1.60	0.93	0.39
OC	1.02	0.60	0.29	1.00	0.59	0.25
S+O	2.60	1.52	0.64	2.60	1.52	0.64

Theory scaled to same (S+O) values and computed for single assumption about the degree of mode decoupling. See also Table 1.

Even without this effect, however, decoupling represented by $x = 0.23$ could explain 98, 98 and 88% of the observed echo power for Europa, Ganymede and Callisto, respectively, where the last number could be increased to 92% by invoking less coupling. These numbers are based on the assumption that the measurements are exact; as they are not, I conclude that essentially all of the echo power from these moons could be due to mode decoupling during retrorefraction. This theory also appears to be consistent with the anomalous dependence of the echoes on wavelength and on the angle of incidence of rays at the surface⁶.

It is not clear whether the experiments to be performed at Jupiter and its moons by the US Galileo orbiting spacecraft will be able to contribute to a more complete understanding of this problem. However, it is evident that new insight would be obtained from bistatic radar studies with separated transmitter and receiver, so that scattering at all values of α instead of only $\alpha = \pi$ could be analysed. In principle, this could be done by using Galileo as a transmitter for receivers on Earth, but estimates of the echo power (R. A. Simpson, personal communication) make the experiment appear marginal. A key requirement for a variety of future planetary radar investigations is the capability of conducting bistatic radar experiments based on powerful transmissions from Earth with reception at spacecraft near the target⁹.

I thank Len Tyler, Dick Simpson, Essam Marouf, Joe Goodman, Tom McCord, Ray Reynolds and Dave Stevenson for helpful discussions. This work was supported by NASA.

Received 10 October 1985; accepted 7 January 1986.

- Campbell, D. B., Chandler, J. F., Pettengill, G. H. & Shapiro, I. I. *Icarus* **34**, 254–267 (1978).
- Goldstein, R. M. & Green, R. R. *Science* **207**, 179–180 (1980).
- Ostro, S. J., Campbell, D. B., Pettengill, G. H. & Shapiro, I. I. *Icarus* **44**, 431–440 (1980).
- Ostro, S. J. & Pettengill, G. H. *Icarus* **34**, 268–279 (1978).
- Hagfors, T., Gold, T. & Ierik, H. M. *Nature* **315**, 637–640 (1985).
- Ostro, S. J. in *Satellites of Jupiter* (ed. Morrison, D.) 213–236 (Arizona University Press, 1982).
- Eshleman, V. R., Tyler, G. L. & Freeman, W. T. *Icarus* **37**, 612–626 (1979).
- Eshleman, V. R. *Science* **205**, 1133–1135 (1979).
- Eshleman, V. R., Tyler, G. L., Simpson, R. A. & Marouf, E. A. *Science* **223**, 772 (1984).

Thermonuclear neutron yield of 10^{12} achieved with Gekko XII green laser

C. Yamanaka & S. Nakai

Institute of Laser Engineering, Osaka University, Suita, Osaka 565, Japan

One of the means by which hydrogen nuclei can be fused, liberating energy in a controllable fashion, involves the implosion of small hydrogen-filled pellets driven by laser radiation. Here we report a world record for thermonuclear neutron yield 1.25×10^{12} per shot, achieved by the Gekko XII green laser system. As the green laser can suppress hot electrons to prevent the preheating of a pellet, high-aspect ratio targets are ablatively imploded with 12 green beams. The multiplexing of shock waves is generated at the centre of a pellet with an accelerating pusher by a quasi-tailored laser pulse. By this scheme the target pusher is uniformly accelerated up to a velocity of almost 10^8 cm s^{-1} without suffering breakup of the target. The temperature of a core is 7 keV, which supports a thermonuclear origin of the neutron yield.

The Gekko XII glass laser system (Fig. 1) consists of 12 beams with total output energy 30 kJ in 1 ns and power 50 TW in 100 ps for a red beam (wavelength $1.05 \mu\text{m}$)¹. The laser was converted to produce a green beam in 1984 and has been used for inertial confinement fusion research.

The direct-driven implosion is divided into three modes: exploding pusher; ablative compression; and shock multiplexing (discussed here). In the case of red-laser irradiation at high intensity, a significant fraction of the absorbed energy is spent in producing hot electrons by resonance absorption^{2–4}. The hot electrons then explode the pusher shell and preheat the fuel, preventing compression—the exploding-pusher mode.

In contrast, the fraction of hot electrons is reduced significantly by using a shorter ($0.53 \mu\text{m}$) wavelength laser. Therefore, even if the target shell is relatively thin, the shell does not explode and the implosion becomes ablative. These considerations encouraged us to try a new type of implosion, shock multiplexing, using a thin-walled shell of large diameter. In this mode the pusher is accelerated for a long distance up to a high velocity resulting in a large neutron yield. All shock waves generated at different times by the accelerated pusher collapse simultaneously^{5,6} at the centre of the pellet. We conclude that the shock-multiplexing mode with thin shelled targets of high aspect ratio is suitable for very high neutron yields ($> 10^{12}$ per shot).



Fig. 1 Main amplifier chain of Gekko XII glass laser system; 12-beam 50 TW, 30 kJ. The laser is converted to green (wavelength $0.53 \mu\text{m}$) by frequency converters settled at the entrance of the final focusing optics.

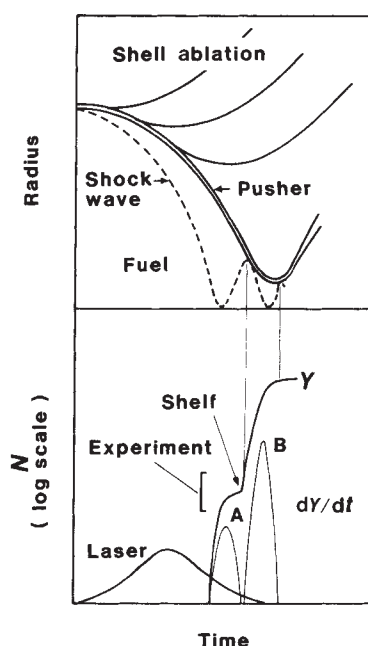


Fig. 2 Schematics of the nominal implosion mode. Outer surface of the fuel container is ablated with intense laser irradiation, and the remainder (pusher) is imploded to compress the fuel. A shock wave is generated by pusher motion, which traverses the fuel region with heating it. Neutron yield predicted by a computer simulation has two plateaux corresponding to the shock-wave reflection A and the adiabatic compression B. Experimentally obtained neutrons were close to the values at the shelf.

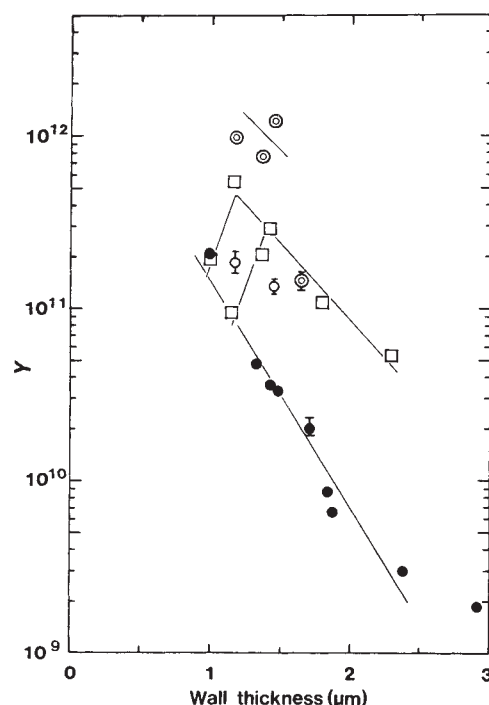


Fig. 4 Dependence of neutron yield on the target-wall thickness. The higher yield is obtained as a consequence of ablatively driven implosion with decreasing wall thickness; the larger the target size, the higher the yield. This is caused by the difference in acceleration time, supported by the observation that the implosion velocity becomes faster for larger aspect ratio targets. The data of 950 μm /750 ps shots infer diameter-pulse matching for a high yield of neutrons. $\odot$, 750 ps/ $\sim$ 950 μm ; $\circ$, 500 ps/ $\sim$ 950 μm ; $\square$, 500 ps/ $\sim$ 730 μm ; $\bullet$, 500 ps/ $\sim$ 500 μm .

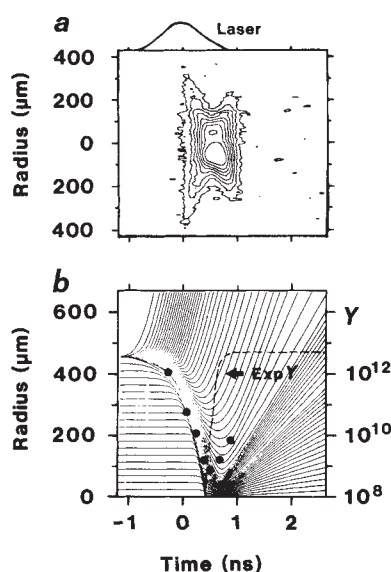


Fig. 3 Implosion dynamics of the highest neutron-yield shot. *a*, Iso-intensity contours of X-ray streaked image; *b*, flow diagram and the history of neutron production replicated by the simulation. Closed circles, peak intensities of the streaked X-ray image; dashed line, simulated neutron yield. The time fiducial of the laser peak is determined by independent measurements with an accuracy of ± 70 ps. The shock waves generated at different times converge at the same point, thus many neutrons are generated in a short time without shelf structure. The observed neutron yield and the implosion behaviour agree with the simulation, where flux limiter (f) = 0.03 is used for electron transport and the absorption is calculated by ray tracing for laser light.

General features of a relatively low aspect ratio thick-shell target implosion are illustrated in Fig. 2. The pusher is accelerated by a reaction of the shell ablation, where a shock wave is generated and travels faster than the pusher. The fuel is heated by the shock wave at the first stage and by the final compression at the second stage. The simulation of the neutron yield with time has a shelf between the first shock reflection and the imploded pusher stagnation. The fusion reaction rate, dY/dt (where Y is neutron yield and t is time), has two peaks in the shock-reflection phase A and in the adiabatic-compression phase B. The neutrons observed in all identical shots are close to the value at the shelf (Fig. 2), which implies that the adiabatic compression of the fuel by the shell stagnation does not contribute to the neutron production. We suspect that the pusher mixes with the fuel during the stagnation, as has been suggested elsewhere^{7,8}. According to theoretical analyses^{7,8}, the high-mode number distortions of the pusher-fuel contact surface grow explosively because of the Rayleigh-Taylor instability in the deceleration phase, and degrade the neutron production. These considerations led us to propose that, to obtain the high-neutron yield, an ablative implosion should be performed without a stagnation phase. An important issue of this mode is the pulse-tailoring, which is necessary to obtain the shock-wave multiplexing. For this purpose we use only the front half-side of a gaussian pulse to select the duration and the timing suitable for the configuration of the pellet.

In these experiments, the laser outputs at 0.53 μm were 7–13 TW in 500 ps and 6–10 TW in 750 ps. The targets were deuterium-tritium filled (3–13 atm) glass micro-balloons of 500–950 μm diameter with aspect ratios of 110–420. The conditions of laser irradiation and target configuration were carefully selected to achieve high uniformity. Figure 3*a* shows iso-intensity contours of streaked X-ray image of the highest neutron shot, and Fig. 3*b* shows the flow diagram of the same implosion

replicated by the one-dimensional simulation code HIMICO^{9,10}. Comparison between Figs 2 and 3 gives a comprehensive view of two different modes.

Figure 4 shows how the neutron yields depend on the thickness of the target pusher. A higher yield is achieved as a consequence of effective acceleration on decreasing the thickness. With an increasing target diameter, a higher yield is expected. However, the neutron yields for the 950 $\mu\text{m}/500$ ps data are lower by an order of magnitude than those for the 950 $\mu\text{m}/750$ ps (Fig. 4), suggesting that the 500-ps pulse is too short for the shock-wave multiplexing in 950- μm targets. The important requirement for shock multiplexing is a continuous increase of pusher acceleration, which is determined by the target configuration and the laser intensity. Simulations allow us to predict an optimized condition. The best matching occurred at the combination of 950 $\mu\text{m}/750$ ps. We detected 1.25×10^{12} neutrons, whose energy corresponds to $\sim 10^{-3}$ of the absorbed energy.

The space-resolved X-ray spectroscopy measurements and the X-ray pinhole images conclusively show that the core is round (30% sphericity) and that the pusher does not mix with the fuel. Plausible mechanisms for stabilization may be: a decrease of the target aspect ratio during the implosion; and

that of the Atwood number on pusher-fuel contact surface. These are caused by increase of the pusher areal density by implosion. Possibility of the impact fusion is excluded by the fuel-temperature measurement. The temperatures deduced from the neutron time-of-flight measurements were 3–7 keV and increased with increasing target-aspect ratio.

We thank K. Mima, T. Yabe and the members of the experimental and theoretical groups at the Institute of Laser Engineering, Osaka University, who contributed to the physical arguments and experimental data described here.

Received 22 November 1985; accepted 16 January 1986.

1. Yamanaka, C. *et al.* 10th int. Conf. Plasma Phys. Controlled Nucl. Fusion Res. (IAEA-BI-1, London, 1985).
2. Manes, K. R. *et al.* *Phys. Rev. Lett.* **39**, 281–284 (1977).
3. Godwin, R. P. *et al.* *Phys. Rev. Lett.* **39**, 1198–1201 (1977).
4. Kruer, W. L. in *Progress in Lasers and Laser Fusion* (eds Kursunoglu, B., Perlmutter, A. & Widmayer, S. M.) **5** (Plenum, New York, 1975).
5. Kidder, R. E. *Nucl. Fusion* **14**, 53–60 (1974).
6. Bertke, S. D. & Goldman, E. B. *Nucl. Fusion* **18**, 509–518 (1978).
7. Freeman, J. R., Clauser, M. J. & Thompson, S. L. *Nucl. Fusion* **17**, 223–230 (1977).
8. Hattori, F., Takabe, H. & Mima, K. *Phys. Fluids* (submitted).
9. Yabe, T. *et al.* *Nucl. Fusion* **21**, 803–816 (1981).
10. Kiyokawa, S. *et al.* *J. appl. Phys. Japan* **22**, L772–L774 (1983).

Observations of solitary waves in a viscously deformable pipe

David R. Scott*, David J. Stevenson†
& John A. Whitehead Jr‡

* Seismological Laboratory and † Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

‡ Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

We have made simple observations of the ascent of a buoyant fluid through a pipe formed in a denser and more viscous fluid that can deform viscously and allow the pipe radius to change. There is no wall between the two fluids, and the Reynolds number is small in both fluids. If the buoyant fluid is supplied at a uniform rate, the system exhibits uniform Poiseuille flow. The response of the system to fluctuations in the rate of supply of the buoyant fluid is to form local maxima in the pipe radius that ascend as solitary waves. Larger-amplitude waves can catch up and collide with smaller waves and, to a good approximation, both waves recover their original form and amplitude after such a collision. Periodic wavetrains are formed when the supply of fluid to the pipe is increased and sustained at a higher rate. These observations gain significance because the system is analogous to that of one-dimensional buoyancy-driven porous flow in a viscous matrix. The experiment may be regarded as a laboratory analogue for studying some aspects of the equations governing porous flow. The observed behaviour is consistent with recent theoretical and computational studies^{1,2}, which have focused on the problem of magma migration in the Earth. The behaviour we observe will, however, arise in other systems governed by the same mechanics. The existence of solitary waves in such systems means that the responses to transient changes in the porosity or the supply of fluid could be long-lived.

Our interest in this system arises from recent advances³ in the understanding of the segregation of magma from a porous matrix of residual crystals, an important process in the crust and mantle of the Earth. Observations of the texture of partially molten rocks^{4,5} have shown that the network of melt may remain interconnected down to very small volume fractions of melt. Recent work identifies the importance of compaction: the deformation of the porous matrix through which the buoyant melt rises. This work, summarized by McKenzie⁶, has prompted the formulation

of a system of equations governing the movement of a buoyant fluid through a porous, deformable matrix; the latter property is essential if magma is to escape from a closed region of partial melt. Our experiment reproduces the essential mechanics of the more complex porous flow problem, and is of interest to those studying porous systems other than partially molten rock.

The experimental apparatus is very simple and can be reproduced easily. A transparent tank, $\sim 10 \text{ cm} \times 10 \text{ cm} \times 50 \text{ cm}$ high, contains transparent viscous fluid. This fluid forms the walls of the deformable pipe. To form the pipe, a buoyant fluid that is less viscous is introduced through a nozzle in the centre of the base of the tank. This fluid is supplied from a reservoir with a variable-pressure head, or by a low-rate pump such as a syringe driver. The buoyant fluid is dyed to make the pipe clearly visible. The experimental fluids we have used are different mixtures of clear honey and water. Diluting honey with water produces a mixture that is less dense and less viscous than pure honey. In Figs 1, 2, the denser fluid in the tank is 100% honey (density 1.42 g cm^{-3} , viscosity 80 poise), and the buoyant fluid in the pipe is 77% honey (density 1.31 g cm^{-3} , viscosity 0.6 poise). The use of mutually-soluble fluids eliminates the effects of surface tension. We have determined that diffusion of water between the two fluids is negligible in the experiment. The use of different mixtures of corn syrup and water also provides good results.

The first part of the buoyant fluid to be introduced ascends as a diapir, but as more fluid is supplied a pipe is established between the nozzle and the top of the tank⁷. If the pressure head of the reservoir of buoyant fluid is held constant, a state of steady Poiseuille flow is established. In this state, the pressure head almost balances the resistance in the tubing supplying the nozzle, because the Reynolds number is small, so there is no pressure difference between the two fluids in the tank. Within the pipe, the buoyancy force balances the non-hydrostatic pressure gradient. We typically use pipe diameters of a millimetre or two.

To create solitary waves, the rate of supply of fluid to the stable pipe must be increased. This can be achieved by increasing the pressure head of the reservoir, or by increasing the supply rate from the syringe driver. In our experiments, the pipe has invariably responded to the change in supply rate by forming solitary waves or wavetrains.

If the increased supply is temporary, followed by a return to the original rate, solitary waves are formed. With care, it is possible to form a single solitary wave. The solitary wave ascends along the uniform pipe at a constant velocity, with unchanging

form. The pipe returns to its original uniform diameter after the passage of the wave. Although the waves are shape-preserving, they are dispersive in that they have an amplitude-dependent phase velocity. A consequence of this is that larger waves catch up and collide with smaller waves; remarkably, both waves survive such a collision almost unscathed (see Fig. 1).

If the increased supply is sustained, a periodic train of solitary waves is formed. This is best demonstrated using a syringe driver, because the rate of supply can be controlled precisely. Figure 2 shows the results of such an experiment. The response to a simple step between two uniform supply rates is a long-lived wavetrain. This differs from the varicose instability⁸, which occurs when the Reynolds number in the interior fluid exceeds a value⁹ of ~ 8 . In this experiment, the maximum Reynolds number is about 3, and the formation of wavetrains is readily observed at smaller Reynolds numbers.

The theoretical model that we have used to compare with the experiments describes the behaviour of fluctuations in the diameter of a single, uniform vertical pipe in an infinite medium. We refer to the viscous wall fluid as the solid phase, and the fluid in the pipe as the liquid. The rigorous description of this system appears to be analytically intractable, but a good approximation can be obtained as follows. As the Reynolds number is small, both in the solid and the liquid, the velocity field satisfies the biharmonic equation and the pressure field satisfies Laplace's equation. If the walls of the tank are distant from the pipe, then the instantaneous pressure distribution in the solid can be represented as a sum of Fourier components of the form $K_0(kr) \sin kz$, where z is the vertical coordinate, r is the radial coordinate, k is the vertical wavenumber, and K_0 is the modified Bessel function of the second kind and zeroth order. We then find that a Fourier component of the radial flow in the solid around the pipe has, in the limit $kr \ll 1$, the form

$$v_r \propto \frac{1}{r} + O(kr^2 \ln kr) \quad (1)$$

We note that we seek *a priori* solitary wave solutions with some wavelength λ . In the limit where the solid viscosity η^s is much greater than the liquid viscosity η^l , we find that $2\pi a/\lambda \sim ka \ll 1$, where $a = a(z)$ is the pipe radius. The stress boundary condition at the pipe wall is then simple, since we need only take the $1/r$ term in the expansion of v_r above; the Fourier sum is trivial and the balance of normal stress across the pipe wall is approximated by

$$p^l = p(z=0) - \rho^s gz + 2\eta^s \frac{v_r(a)}{a} \quad (2)$$

where ρ^s is the solid density, g is the acceleration due to gravity and v_r is now the total radial flow at $r = a$. The Poiseuille flow formula for the flow in the pipe, neglecting the slope of the walls, is

$$u = -\frac{\pi a^4}{8\eta^l} \left[\frac{\partial \rho^l}{\partial z} + \rho^l g \right] \quad (3)$$

where u is the volume flux of liquid through the pipe relative to the solid and ρ^l is the liquid density. We link the flow in the solid and liquid using continuity. Clearly, $v_r(a) = \partial a / \partial t$, where t is time. Also,

$$\frac{\partial u}{\partial z} = -\frac{\partial}{\partial t} (\pi a^2) \quad (4)$$

Substituting the z -derivative of equation (2) into equation (3) and using these continuity conditions, we obtain

$$u = \frac{\pi a^4}{8\eta^l} \left[g(\rho^s - \rho^l) + \eta^s \frac{\partial}{\partial z} \left(\frac{1}{\pi a^2} \frac{\partial u}{\partial z} \right) \right] \quad (5)$$

As expected, equations (4) and (5) are analogous to those governing porous flow in a deformable matrix, when the flow

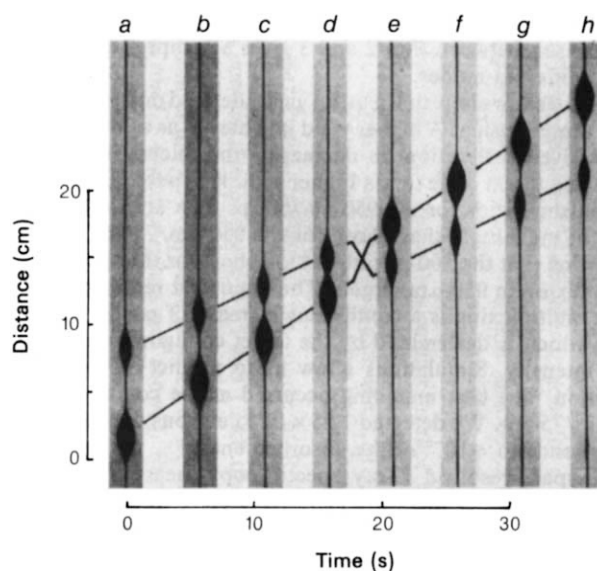


Fig. 1 The interaction of two solitary waves. The photographs were taken using back lighting and a flat-sided tank, to avoid distortion. Only part of the tank is shown in each frame; the tank is 10-cm wide. The frames are spaced according to their separation in time. In frames *a-c* and *f-h*, the two waves ascend as solitary waves with constant amplitude, shape and velocity. This is shown by the ruled lines. The main outcome of the interaction seen in frames *d, e* is the phase shift, shown by the offset of the ruled lines. The amplitudes of the two waves are conserved very well; our measurements show that the larger waves gain a little liquid at the expense of the smaller waves. The uniform pipe is undisturbed by either the passage of the waves or the interaction.

is constrained to one dimension (vertical). This similarity persists even if the tank walls are closer to the conduit and 'plug flow' occurs in the solid¹⁰. The pipe equations are obtained by substituting the porosity $f \approx \pi a^2$, choosing an f^2 dependence for the permeability, and choosing a $1/f$ dependence for the bulk viscosity of the solid matrix.

Some analysis, and numerical modelling, of solitary wave solutions to the one-dimensional porous flow equations has already been presented^{1,11}. In particular, Scott and Stevenson¹ give explicit expressions for the dispersion relation between wave amplitude and phase velocity. For the special case of a deformable pipe, this relationship is:

$$c = \frac{a_0^2 g (\rho^s - \rho^l)}{8\eta^l} 2 \frac{\left[\ln \Psi - \frac{1}{2} + \frac{1}{2\Psi^2} \right]}{\left[1 - \frac{2}{\Psi} + \frac{1}{\Psi^2} \right]} \quad (6)$$

where a_0 is the radius of the uniform part of the pipe, c is the phase velocity and $\Psi = (a_{\max}/a_0)^2$. The leading quotient on the right-hand side is the mean Poiseuille flow velocity in the uniform pipe.

Using photographs of the experiment, we have made limited measurements of the amplitudes, ascent velocities, and wavelengths of the waves; these agree with the theory to within 10%. Waves with values of Ψ between 6 and 80 were used. Thus, we find good agreement between measurements of this experiment, measurements of numerical experiments¹, and the theoretical dispersion relation. The experiments presented in Figs 1, 2 complement the numerical experiments presented earlier¹, where we similarly showed a collision and the formation of a wavetrain. This correspondence lends credibility to the numerical studies. We note that more detailed measurements have been carried out by Olson and Christensen¹².

Real porous media have a much more complicated micro-

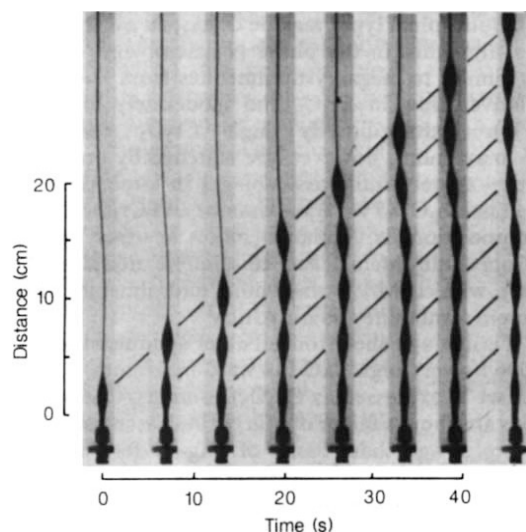


Fig. 2 The formation of a periodic wavetrain. This experiment started with a uniform pipe, sustained by a constant supply of buoyant liquid from the nozzle seen at the base of the frames. The rate of supply was then increased and held at a new constant level, approximately 220 times higher. This increase could be accommodated by the formation of a larger uniform pipe, but the system preferentially forms a periodic wavetrain. In longer experiments, the wavetrain does slowly degrade down to a larger uniform pipe. The ruled lines are parallel and equally spaced.

scopic structure than that discussed here; this will change the functional form of the matrix permeability and viscosity, but not the existence and nature of the waves described here. More importantly, real porous media are three-dimensional; numerical experiments¹⁰ indicate that the one-dimensional waves are unstable in higher dimensions, but that solitary waves of higher dimensions are formed. The balance of dissipative mechanisms and gravitational energy release governing the ascent of these waves is the same in one, two, or three dimensions. We believe, therefore, that the experiments described here are not only pedagogically useful, but also serve as a laboratory analogue for real Earth processes. Specifically, we may expect the geochemistry, morphology, and timing of igneous processes to be affected by phenomena resembling those reported here: this is discussed elsewhere¹⁰. It is possible that solitary waves arise in other porous systems.

We thank Robert E. Frazel at Woods Hole, and various members of the Division of Geological and Planetary Sciences at Caltech, for experimental assistance. Peter Olson and Uli Christensen have independently performed experiments similar to those described here, and have shared their results. This work was supported by NSF grant EAR-8418353, and by the Geodynamics program of the Center for the Analysis of Marine Systems at Woods Hole Oceanographic Institution. Contribution number 4271, Division of Geological and Planetary Sciences, California Institute of Technology.

Received 29 September; accepted 30 December 1985.

1. Scott, D. R. & Stevenson, D. J. *Geophys. Res. Lett.* **11**, 1161-1164 (1984).
2. Richter, F. M. & McKenzie, D. P. *J. Geol.* **92**, 729-740 (1984).
3. McKenzie, D. P. *J. Petrol.* **25**, 713-765 (1984).
4. Wafi, H. S. & Bulau, J. R. *J. geophys. Res.* **84**, 6109-6114 (1979).
5. Cooper, R. F. & Kohlstedt, D. L. *Tectonophysics* **107**, 207-233 (1984).
6. McKenzie, D. P. *Earth planet. Sci. Lett.* **74**, 81-91 (1985).
7. Whitehead, J. A. & Luther, D. S. *J. geophys. Res.* **90**, 705-717 (1985).
8. Whitehead, J. A. *Geophys. J. R. astr. Soc.* **70**, 415-433 (1982).
9. Huppert, H. E., Sparks, R. S. J., Whitehead, J. A. & Hallworth, M. A. *J. geophys. Res.* (in the press).
10. Scott, D. R. & Stevenson, D. J. *J. geophys. Res.* (in the press).
11. Barclon, V. & Richter, F. M. *J. Fluid Mech.* (in the press).
12. Olson, P. & Christensen, U. *J. geophys. Res.* (submitted).

Metasomatic mineral titanate complexing in the upper mantle

Stephen E. Haggerty*, A. J. Erlank† & Ian E. Grey‡

* Department of Geology, University of Massachusetts, Amherst, Massachusetts 01003, USA

† Department of Geochemistry, University of Cape Town, Rondebosch 7700, South Africa

‡ CSIRO Division of Mineral Chemistry, PO Box 124, Port Melbourne, Victoria, Australia 3207

The extent and expression of metasomatism in the upper mantle are contentious issues^{1,2}, although the process has gained widespread acceptance to account for the subsequent enrichment of previously depleted lithosphere^{3,4}. Uncertainties exist in the origin and precise compositions of the fluids involved, as well as in the total inventory of introduced metasomatic elements (particularly silicate-incompatible Ti, Zr, Nb, REE (rare-earth elements), Ba, Sr, Rb, K and Na). Metasomatism is readily identified by the presence of phlogopite, and in the most advanced stages by K-richterite⁴. In K-richterite-bearing harzburgites, these characteristic minerals are accompanied by rutile, armalcolite and lindsleyite-mathiasite (LIMA) solid solution members⁵ of the crichtonite mineral series⁶. The LIMA series is one upper mantle repository for silicate-incompatible elements, but the series is also specifically enriched in chromium (12-18 wt% Cr₂O₃). While it has been assumed^{5,7} that Cr is immobilized in the depleted lithosphere as a restite element, either in residual garnet, pyroxene, or in chromian spinel, there has been no direct evidence for the source of Cr in LIMA. We show here that LIMA solid-solution members are derived from Cr-spinel through interaction with metasomatic fluids by formation of a magnetoplumbite mineral with up to 14 wt % BaO, and which is inferred to contain Ce⁴⁺. Fluid compositions are heterogeneous and an increase in the variety of repositories for silicate-incompatible trace elements is demonstrated in the subcontinental lithosphere at depths <100 km.

Lindsleyite (Ba-specific) and mathiasite (K-specific) are two new minerals⁵ in the crichtonite mineral series, AM₂₁O₃₈ (where A = large cations Ba, K, Sr, Pb, Na, Ca, REE and M = small cations Ti, Cr, Nb, Fe, Mg, Al, Mn, Zn, Zr) that develop in upper-mantle depleted harzburgites of subcontinental lithospheric origin. In their description of the new minerals, S.E.H. *et al.*⁵ illustrated (their Fig. 1b) an assemblage of chromian spinel cores rimmed by lindsleyite and this assemblage prompted a more detailed study of spinel-forming reactions during metasomatism. Samples examined were collected from the waste dumps (Bultfontein Floors) of country rock and boulder-sized upper mantle xenolith concentrates that have resulted from early diamond mining in the Kimberley district of South Africa. One sample (BD-3096) is a complexly veined harzburgite. The substrate is dominated by coarse (~5 mm) equant crystals of olivine and orthopyroxene, and subordinate 1-3-mm laths of phlogopite and anhedral K-richterite. The two largest veins (3-5-mm width) are composed of K-richterite, with lesser and corroded olivine, diopside, phlogopite and opaque mineral oxides. Euhedral to subhedral lindsleyite (Fig. 1a) and rounded grains of chromian spinel are present as rare inclusions in phlogopite and K-richterite, both in substrate and in veins. Interstitial to olivine and orthopyroxene, but always closely associated with phlogopite and K-richterite, are abundant, irregularly shaped assemblages of opaque mineral oxides that are commonly amoeboidal and 1-3 mm in the longest dimension.

Five distinct minerals are identified in the opaque oxide mineral assemblage using high-resolution oil-immersion objectives and reflected light microscopy. Approximately equal proportions of chromian spinel and an unidentified mineral (phase N), rimmed by lindsleyite, coexist in a blocky fabric or in an

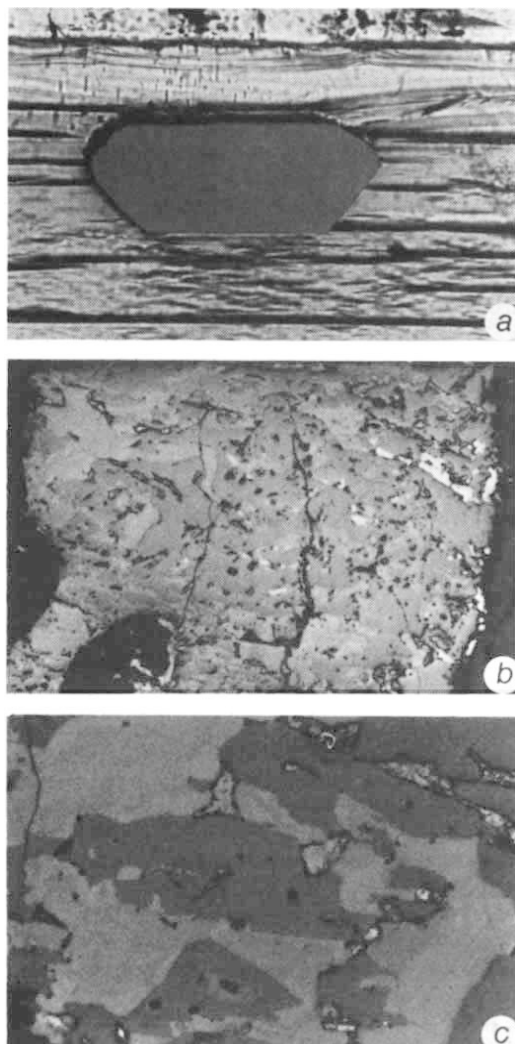


Fig. 1 Reflected light photomicrographs of *a*, euhedral equilibrated lindsleyite enclosed in phlogopite (air objective, field width 0.48 mm); *b*, a blocky and symplectic five-phase complex composed of dark-grey Cr-spinel, lighter-grey phase N, fringed by thin borders of lindsleyite, white Nb-Cr rutile, and intermediate-grey Cr-picroilmenite (air objective, field width 0.48 mm); *c*, a higher magnification of the left-hand side of the complex shown in *b* illustrating the presence of lindsleyite as an annulus around phase N and Cr-spinel (glycerine objective, field width 0.12 mm). The textural interpretation and conclusion reached is that Cr-spinel has reacted with metasomatizing fluids, has partially decomposed and has given rise to four additional minerals, all of which are highly enriched in silicate-incompatible elements.

intricate symplectite (Fig. 1*b*), with accessory ilmenite and rutile dispersed throughout the intergrowth (Fig. 1*b*).

Electron microbeam analyses of the mineral assemblages show that lindsleyites enclosed in phlogopite (Fig. 1*a*) or K-richterite are compositionally very similar (Table 1), both to each other and to lindsleyites reported^{5,8} in other polished thin sections of BD-3096. A comparison of these lindsleyites, however, with lindsleyite in the Cr-spinel, phase N, ilmenite and rutile intergrowth (Fig. 1*b*) shows that phase-N-associated lindsleyite has higher TiO₂ (by ~3 wt %) and BaO (5.9 compared with 5.1 wt %) concentrations, significantly lower ZrO₂ contents (0.3 compared with 3.8 wt %), and approximately equal proportions of Cr₂O₃, FeO, MgO, SrO, K₂O, Nb₂O₅, La₂O₃ and Ce₂O₃. Spinel in the phase N assemblage is enriched in TiO₂ relative to spinel enclosed in phlogopite, but has lower contents of R³⁺ ions, higher concentrations of MgO (9.0 compared with 6.5 wt %), but virtually identical average compositions of FeO

(Table 1). Both spinel types may be described as Mg-Ti-ferrian chromites. Ilmenites in the phase N assemblages are compositionally similar to megacrystic ilmenites from kimberlites^{9,10}, being relatively low in Fe₂O₃ and moderately high in MgO (~10 wt %). Extraordinarily high Cr₂O₃ concentrations (9.5 wt % maximum), however, are matched by only one suite of ilmenite + spinel + rutile assemblages in kimberlite¹¹. Rutile is enriched in Nb₂O₅ (3 wt % maximum) and Cr₂O₃ (5–7 wt %), and is compositionally similar to rutiles in other veined peridotites from Bultfontein⁷, and to discrete nodules of rutile intergrown with LIMA, armalcolite and ilmenite from the Jagersfontein kimberlite, South Africa¹².

Phase N is distinguished from all other opaque mineral oxides in BD-3096 by very high BaO (14 wt % maximum) and cerium oxide (3.1 wt % expressed as Ce₂O₃) contents. Chromium concentrations are about a factor of 3 larger than associated lindsleyite, but are lower than those of Mg-Ti-ferrian chromite. Titanium is an essential component, with a range of 22–24 wt % TiO₂. Average magnesium contents (~3.5 wt % MgO) of phase N and lindsleyite are identical, as are iron values (expressed as FeO) for phase N and coexisting spinel and ilmenite (~25 wt % FeO). The large composite grain shown in Fig. 1*b* was excavated from the polished thin section and fragmented. Individual fragments were studied in a scanning electron microscope and by X-ray precession techniques. Cell dimensions, symmetry and intensity distribution in the diffraction patterns are consistent with phase N being isostructural with barium ferrite, BaFe₁₂O₁₉ (ref. 13), with the magnetoplumbite structure¹⁴, and its orientation relationship to spinel is $[0001]_N \parallel [111]_{\text{spinel}}$, $(30\bar{3}0)_N \parallel (224)_{\text{spinel}}$. The general formula for magnetoplumbite phases is $AM_{12}O_{19}$, where A = large cations with 12-fold coordination and M = small cations with 4-, 5- and 6-fold coordination to oxygen. If the large cations (Ba, K and Ce³⁺) in phase N are assigned to the A-cation site, the sum of the small cations is only 10.06, far short of the 12 required for the magnetoplumbite composition. In a detailed structural analysis to be published elsewhere, we show that if Ce is considered to be a small cation, it is tetravalent and the formula for phase N is $A_{1.00}M_{11.84}O_{19}$ (Table 1), which is close to the $AM_{12}O_{19}$ magnetoplumbite composition. On this basis, an adjustment of the Fe²⁺/Fe³⁺ ratio is required and an approximate 1:1 ratio results (Table 1).

The five-phase assemblage (Mg-Ti-ferrian chromite, phase N, lindsleyite, chromian-picroilmenite and Nb-Cr rutile) is dominated by Cr₂O₃. Relative to spinel in phlogopite, the assemblage also contains proportionately higher contents of Ba, Sr, K, Ti, Nb, Ce and La. These elements (and also Na in view of its abundance in K-richterite) are regarded unequivocally as being of metasomatic origin. Average iron contents of spinel in phlogopite are similar to spinel, phase N and ilmenite in the five-phase assemblage, implying from mass balance and relative mineral proportions that very little iron was added by the metasomatic media; this also applies to Mg. Zirconium, which is typically an essential component of lindsleyite⁵, is in very low abundance in the phase N assemblage compared with equilibrated lindsleyite enclosed in phlogopite or K-richterite, suggesting that Zr may be decoupled from the package of LIMA-characteristic elements by gradients in the activity of H₂O; high a_{H_2O} , coupled with alkalis, may be more effective for Zr transport, possibly as a (Na, K)₄ZrSi₂O₈ complex similar to that modelled for alkali-rich felsic melts¹⁵. The REE are concentrated in phase N with Ce > La, and decrease proportionately in lindsleyite as a consequence of the dilution by interacting spinel. Niobium is strongly partitioned into rutile, followed by lindsleyite, phase N and ilmenite.

A comparison of the analyses for phase N and associated spinel (Table 1) shows that the main composition changes are due to replacement of Cr and Mg in the spinel by Ti and Ba. The magnetoplumbite structure comprises slabs of spinel-type structure, four anion-layers thick, intergrown with perovskite-

Table 1 Chemical formulae from electron microbeam analyses

(1) Phase N	$(\text{Ba}_{0.90}\text{K}_{0.10})[\text{Cr}_{4.20}\text{Ti}_{2.96}\text{Fe}_{1.80}^{3+}\text{Fe}_{1.71}^{2+}\text{Mg}_{0.88}\text{Ce}_{0.16}(\text{Al}, \text{Zr}, \text{Mn}, \text{Si}, \text{Nb}, \text{Ta})_{0.13}]\text{O}_{19}$
(2) Lindsleyite	$(\text{Ba}_{0.71}\text{Sr}_{0.22}\text{Ca}_{0.05}\text{K}_{0.05})(\text{Ti}_{13.08}\text{Cr}_{2.99}\text{Fe}_{3.37}\text{Mg}_{1.60}\text{Mn}_{0.04}\text{Zr}_{0.04}\text{Nb}_{0.13}\text{La}_{0.06}\text{Ce}_{0.17})\text{O}_{38}$
(3) Spinel	$(\text{Fe}_{5.85}^{2+}\text{Fe}_{1.71}^{3+}\text{Mg}_{3.73}\text{Mn}_{0.10}\text{Nb}_{0.01}\text{Al}_{0.21})(\text{Cr}_{10.73}\text{Ti}_{1.63}\text{Zr}_{0.01})\text{O}_{32}$
(4) Ilmenite	$(\text{Fe}_{2.50}^{2+}\text{Fe}_{0.50}^{3+}\text{Mg}_{0.06}\text{Mn}_{0.01})(\text{Cr}_{0.17}\text{Ti}_{0.88})\text{O}_3$
(5) Rutile	$(\text{Fe}_{0.02}^{2+}\text{Mg}_{0.02}\text{Nb}_{0.02})(\text{Cr}_{0.06}\text{Ti}_{0.91})\text{O}_2$
(6) Lindsleyite	$(\text{Ba}_{0.59}\text{Sr}_{0.29}\text{Ca}_{0.05}\text{K}_{0.11}\text{Na}_{0.01})(\text{Ti}_{12.37}\text{Cr}_{3.07}\text{Fe}_{3.34}\text{Mg}_{1.31}\text{Mn}_{0.01}\text{Zr}_{0.55}\text{Nb}_{0.12}\text{La}_{0.02}\text{Ce}_{0.15})\text{O}_{38}$
(7) Spinel	$(\text{Fe}_{5.49}^{2+}\text{Fe}_{2.69}^{3+}\text{Mg}_{3.01}\text{Mn}_{0.16}\text{Al}_{0.08})(\text{Cr}_{12.02}\text{Ti}_{0.54}\text{Zr}_{0.02})\text{O}_{32}$

Analyses (1)–(5) are average compositions from the complex intergrown assemblage shown in Fig. 1b. Compositions (6) and (7) are averages from analyses of lindsleyite (Fig. 1a) and spinel enclosed in phlogopite or K-richertite.

like close-packed layers parallel to (111)_{spinel}. To a first approximation, the structure of phase N may be considered to derive from spinel by periodic replacement of the octahedron and one tetrahedron from the spinel T₂ cation layer (containing mainly Cr and Mg, respectively) by a segment of hexagonal BaTiO₃. Yiemengite, K(Cr, Ti, Fe, Mg)₁₂O₁₉, is a recently described¹⁶ magnetoplumbite mineral from a kimberlite in the Shandong Province, China, and is associated with chromite and mathiasite. It is, therefore, similar to the assemblage we describe in BD-3096, which is Ba-specific, whereas mathiasite and yiemengite are dominated by K.

Major and minor element chemistries of the opaque mineral oxide assemblages in BD-3096 may be divided into spinel-characteristic and metasomatic-characteristic classes. Chromium and Mg fall into the first group; the second group consists of Ba, Ce, La, Ti, Zr and Nb with minor K, Sr and Ca; and Fe may be present in both classes. Under the effects of metasomatism, reconstitution of Cr-spinel results in structurally coherent phase N and the formation of lindsleyite. Phase N has high Ba and Ce concentrations so that some fraction of the fluid involved was charged with these elements in the presence of Ti. Zirconium is present in high concentrations only in equilibrated lindsleyite enclosed in phlogopite or K-richertite, implying that Zr transport in the metasomatic fluids may be preferentially enhanced by high alkali (Na + K) and water contents. Inferred Ce⁴⁺ in phase N and Fe³⁺/Fe²⁺ ratios in Cr-spinel and ilmenite imply metasomatic redox conditions somewhat higher than those determined for the lower lithosphere¹⁷.

It is estimated that the titanate complex, in association with phlogopite and K-richertite, formed in depleted harzburgite at 900–1,100 °C and pressures of 20–30 kbar (refs 5, 12, 18), placing the active zone of metasomatism at 75–100 km in the subcontinental lithosphere. In a wider context, the metasomatism discussed here has been shown to be unrelated to and older than kimberlite emplacement at 90 Myr (ref. 4), but there is uncertainty as to whether the fluids concerned are derived from MARID (mica, amphibole, rutile, ilmenite, diopside)-type pegmatites related to kimberlite magmatism at depth^{7,19}, or whether they are of different asthenospheric origin^{4,20}. Although such uncertainty clouds the issue of the extent and scale of this metasomatism as an upper mantle process, its style and nature are important in elucidating mantle processes. Thus, the high concentrations of Na, K, Ba, Sr, Ti, Zr, Nb and REE involved in the formation of phlogopite, K-richertite and the exotic opaque mineral assemblage described here attest not only to the possible mineral repositories for silicate-incompatible trace elements in the subcontinental lithosphere, but also to the compositional complexity of the fluids and mineral-forming reactions that have produced these assemblages. That similar minerals and assemblages are found in places as far apart as China and South Africa broadens the scope for considering that assimilation of metasomatic products is primarily responsible for the geochemical signatures imparted to kimberlites and perhaps also lamproites.

This study was supported by the NSF (EAR83-08297 to S.E.H.), the University of Cape Town, CSIR and CSIRO. We

thank J. B. Dawson for supplying BD-3096 and other metasomites from the Bultfontein Floors.

Received 11 December; accepted 30 December 1985.

1. Bailey, D. K. *Nature* **296**, 525–530 (1982).
2. Dawson, J. B. in *Developments of Petrology IIB, Kimberlites: The Mantle and Crust–Mantle Relationships* Vol. 2, 289–294 (Elsevier, Oxford, 1984).
3. Erlank, A. J., Allsopp, H. L., Duncan, A. R. & Bristow, J. W. *Phil. Trans. R. Soc. Lond. A297*, 295–307 (1980).
4. Erlank, A. J., Allsopp, H. L., Hawkesworth, C. J. & Menzies, M. A. *Terra Cognita* **2**, 261–263 (1982).
5. Haggerty, S. E., Smyth, J. R., Erlank, A. J., Rickard, R. S. & Danchin, R. V. *Am. Miner.* **68**, 494–505 (1983).
6. Grey, I. E., Lloyd, D. J. & White, J. S. Jr *Am. Miner.* **61**, 1203–1212 (1976).
7. Jones, A. P., Smith, J. V. & Dawson, J. B. *J. Geol.* **90**, 435–453 (1982).
8. Jones, A. P. & Ekambaram, V. *Am. Miner.* **70**, 414–418 (1985).
9. Mitchell, R. H. *J. Geol.* **81**, 301–314 (1973).
10. Haggerty, S. E. in *Oxide Minerals, Reviews of Mineralogy* Vol. 3 (ed. Rumble, D. III) 101–300 (American Mineralogical Society, 1976).
11. Wyatt, B. A. in *The Mantle Sample: Inclusions in Kimberlite and Other Volcanics* Vol. 1 (eds Boyd, F. R. & Meyer, H. O. A.) 257–264 (American Geophysical Union, Washington DC, 1979).
12. Haggerty, S. E. *Geochim. cosmochim. Acta* **47**, 1833–1854 (1983).
13. Townes, W. D., Fang, J. H. & Perrotta, A. J. *Z. Kristallogr.* **125**, 437–449 (1967).
14. Adelskold, V. *Ark. Kemi Miner. Geol. Ser. A12*, 1–9 (1938).
15. Watson, E. B. *Contr. Miner. Petrol.* **70**, 407–419 (1979).
16. Dong, Z., Jianxiong, Z., Ol, L. & Peng, Z. *Kexue Tongbao (Bul. Sci.)* **15**, 932–936 (Foreign language edn **29**, 920–923; 1983).
17. Haggerty, S. E. & Tompkins, L. A. *Nature* **303**, 295–300 (1983).
18. Podpora, C. & Lindsley, D. H. *Eos* **65**, 293 (1984).
19. Kramers, J. D., Roddick, J. C. M. & Dawson, J. B. *Earth planet. Sci. Lett.* **65**, 90–106 (1983).
20. Hawkesworth, C. J., Erlank, A. J., Marsh, J. S., Menzies, M. A. & van Calsteren, P. in *Continental Basalts and Mantle Xenoliths*, 111–138 (Shiva, Cheshire, 1983).

Chlorobiaceae in Palaeozoic seas revealed by biological markers, isotopes and geology

Roger E. Summons*† & Trevor G. Powell†

* Baas Becking Geobiological Laboratory and † Division of Continental Geology, Bureau of Mineral Resources, Geology and Geophysics, Canberra, ACT, Australia 2600

Biomarker hydrocarbons in oils and sediments have considerable potential for assessing the nature of biological processes in the past, but only a few can be uniquely related to particular organisms. Reef-hosted oils in the Silurian of the Michigan Basin, Canada and the Devonian of Western Canada have high abundances of 1-alkyl-2,3,6-trimethylbenzenes. Their precise structures suggest that they are diagenetic products of aromatic carotenoids of the green sulphur bacteria. Individual components are enriched in ¹³C by 7–8‰ relative to the saturated hydrocarbons of the same oil. An isotopic anomaly of this magnitude and direction correlates with a component of photosynthetic carbon assimilation by the reductive tricarboxylic acid cycle. These results, together with data on palaeoenvironments, provide compelling evidence for the existence, in ancient restricted seas, of microbial communities containing *Chlorobiaceae*, organisms with distinctive biochemistry which leave morphologically identifiable remains.

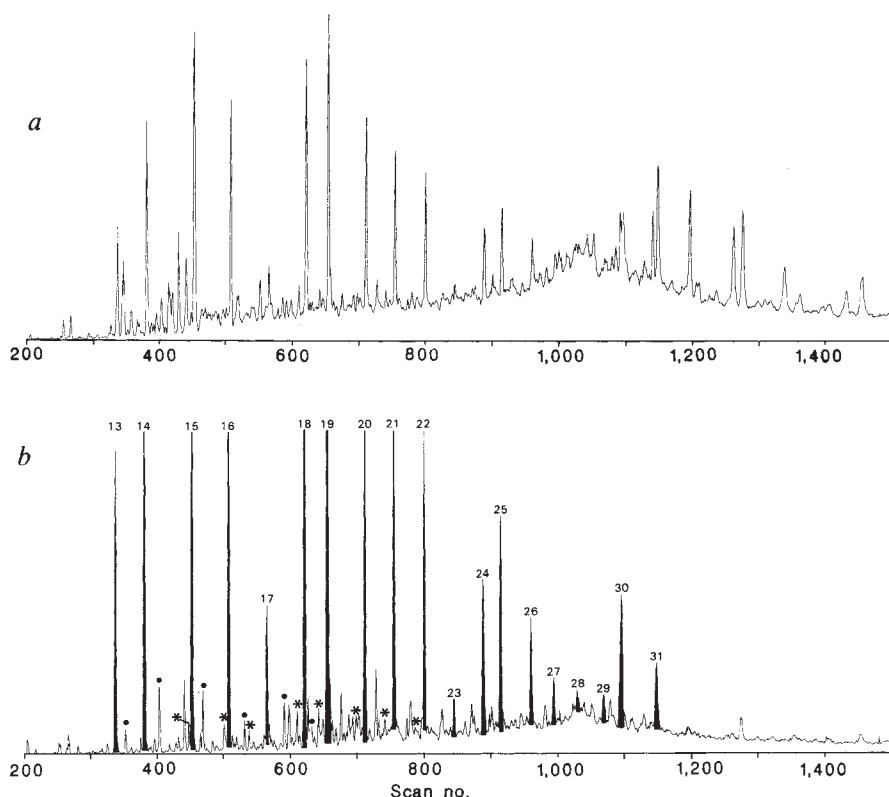


Fig. 1 Reconstructed ion chromatograms from GC/MS analysis of monoaromatic fraction from Ram oil. *a*, Total ion current. *b*, m/z 133. Shaded peaks are 1-alkyl-2,3,6-trimethylbenzenes (I, Fig. 2); peaks marked • and * are minor isomers. Components above scan 1,100 are alkylbenzenes unrelated to the trimethyl compounds.

The presence of distinctive 'biomarker' hydrocarbons is often used to group oils into related families and to assign their origins to specific source rocks¹. Similarly, small differences (1–3‰) in stable carbon isotope ratios between bulk hydrocarbon fractions of different oils can also be used to determine genetic differences^{2,3}. Oils trapped in Silurian pinnacle reef reservoirs in the Michigan Basin (see Ram 15 Eniskillen in Table 1) have unusual concentrations of individual alkyl benzenes, which have been used to correlate the oils to their parent source rocks⁴. Oils (such as Zama) in similar pinnacle reefs in Middle Devonian rocks of the Alberta Basin, Canada⁵ contain the same compounds, but at lower concentrations. The characteristic distribution pattern, in which the C_{17} , C_{23} and C_{28} components are in low abundance (Fig. 1), and the occurrence of an intense benzylic cleavage at m/z 133 in their mass spectra suggests that these compounds are trimethyl-substituted aryl isoprenoids with tail-to-tail irregular linkage in components of >20 carbon atoms.

Aryl isoprenoids of the type shown in Fig. 2 have previously been reported from a Russian oil of unspecified age⁶. Their origin was attributed to disproportionation during diagenesis of a β -carotene-like precursor. As the reported chemical synthesis⁶ aptly demonstrates, such an origin would yield an isomeric mixture dominated by the most thermodynamically stable compounds formed by the necessary migration of a ring methyl group. Examination by capillary gas chromatography/mass spectrometry (GC/MS) reveals that the oils reported here are dominated by a single isomer of each compound with only traces of alternative ring substitution patterns (Fig. 1). Totally unambiguous synthesis of five likely ring isomers, together with comparison of nuclear magnetic resonance (NMR) and GC/MS data with components from the oils, has confirmed that the ring is methyl-substituted at carbons 2, 3 and 6 relative to the isoprenoid chain (unpublished work; see Fig. 2, structure I).

Three C_{40} diaromatic carotenoid-derived hydrocarbons have been identified in Toarcian shales of the Paris Basin⁷ and contain two aromatic ring substitution patterns (2,3,4- and 2,3,6-). These patterns occur singly or together in the diaromatic carotenoids renieratene I, isorenieratene and renierapurpurin found in extant photosynthetic bacteria and sponges⁸ (or their bacterial

symbionts^{9,10}) and in the monoaromatic carotenoids okenone III and chlorobactene IV found in photosynthetic bacteria^{11,12} (Fig. 2). The diaromatic compounds were not detected in the oils reported here. Selective diagenetic hydrogenation and cleavages along the isoprenoid chain are often observed in this class of compound^{7,13}. The greater probability of scission next to branch points (Fig. 2, I) accounts for the distinctive distribution pattern of the mixture (Fig. 1) and the progressive decrease in the relative abundance of the higher homologues. Such a process operating on the C_{40} diaromatic hydrocarbons found in the Paris Basin sediments⁷ would yield products with two aromatic substitution patterns. However, chlorobactene IV and iso-renieratene are the characteristic and relatively abundant pigments of extant green sulphur bacteria *Chlorobiaceae*¹³ and would yield exclusively 2,3,6-methyl substituted isomers by simple chain scission.

The presence of high concentrations of aryl isoprenoids in the Silurian oils is accompanied by a carbon isotope anomaly

Table 1 Isotope data for Ram 15 Eniskillen and Zama oils

Carbon isotope composition* of hydrocarbon fractions		
Ram oil	^{13}C (‰)	Zama oil ^{13}C (‰)
Saturates	−30.7	−30.5
Aromatics	−28.5	−28.41
Resins	−27.8	ND
Asphaltenes	−28.6	−28.69
Carbon isotope composition* of major aryl isoprenoids† in Ram oil		
	^{13}C (‰)	
C_{15}	−22.55	
C_{18}	−22.14	
C_{19}	−21.61	
C_{20}	−24.74	
C_{21}	−22.71	

* Relative to PDB standard; † obtained by repeated collection from preparative GC; ND, not determined.

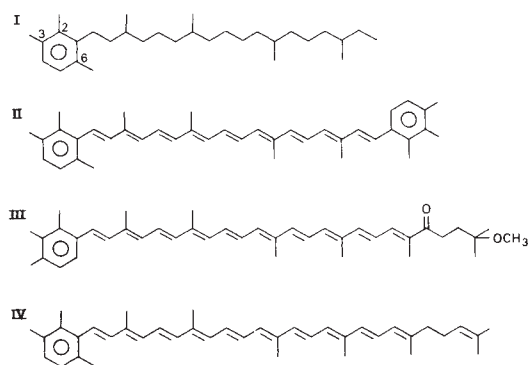


Fig. 2 Fossil and extant bacterial carotenoids: I, Generalized structure of C_{13} to C_{31} hydrocarbons identified in Ram Eniskillen 15 oil; II, renieratene; III, okenone; IV, chlorobactene.

in which the aromatic fractions are up to 3% heavier than the corresponding saturate fractions⁴ (for example, Ram oil, Table 1). This contrasts with the normal situation in marine oils where the isotopic difference between saturates and aromatics in oils is <1.3% (ref. 3). The Zama oil from northern Alberta shows a similar but less marked anomaly (Table 1). Moreover, the heavy isotope signature of the Ram oil aromatics can be attributed to the relatively large concentration of the few aryl isoprenoids enriched in ^{13}C by 7–8% compared with the saturate fraction (Table 1). This enrichment in ^{13}C assumes additional relevance as it is tied to a group of individual compounds and it is in the opposite sense to the fractionation normally observed for lipids relative to the total biomass¹⁴.

Oil-source correlation studies have shown that the source of both sets of oils was organic-rich, inter-reef sediments which were deposited under metahaline to hypersaline sulphate and sulphide-rich water columns^{4,5}. Both source rock sequences consist of dolomites which interfinger with overlying sulphate-evaporites and reflect progressive restriction of the overlying water body. The extremely low pristane-to-phytane ratios (<0.5) in the saturate fractions from these source rock sequences^{4,5} attest to strongly reducing conditions. Such environments are favoured by anoxygenic photosynthetic sulphur bacteria (*Chromatiaceae* and *Chlorobiaceae*)^{15,16}. *Chlorobium* species generally underlie the motile *Chromatium* species and its complementary pigment composition allows it to harvest light not used by *Chromatium*¹⁶. The latter, operating as photoautotrophs, fix carbon through the Benson-Calvin pathway, whereas *Chlorobium* uses a reverse or 'reductive' Krebs' cycle¹⁷. An important consequence of this difference is that organic matter produced by these organisms and using a typical marine bicarbonate source ($^{13}C_{PDB} = 0\%$) will have markedly different carbon isotope compositions of –26 to –36‰ for *Chromatium* and –9 to –21‰ for *Chlorobium*^{18–20}. Other processes leading to abnormally heavy carbon are photosynthesis in C_4 and crassulacean acid metabolism (CAM) terrestrial plants and photosynthesis where bicarbonate is a limiting substrate, possibly including situations in extremely evaporitic environments^{20,21}. There is no evidence for the contribution of terrestrial organic matter to the source environment^{4,5} and the latter situation would lead to ^{13}C enrichment of the total organic matter and not merely the aromatic isoprenoid precursors. The isotopic anomaly is largely attributable to the aryl isoprenoids as they are derived simply from specific carotenoids from the *Chlorobiaceae*, which have an unusual carbon metabolism. These carotenoid hydrocarbons would be relatively inert to decay, leading to their preferential preservation compared with oxygenated compounds. The otherwise diverse biota of the source environment is indicated in the oils by the presence of the usual linear, branched and cyclic alkanes and suites of C_{27} – C_{30} steranes and C_{27} – C_{35} triterpenoid biomarkers.

These results are a unique correlation of the identification of chemical fossils with a specific isotope signature. This, together with geological evidence of palaeoenvironments, allows the origin of these hydrocarbons to be assigned with a high degree of confidence, and provides a record for the occurrence of *Chlorobiaceae* in two ancient restricted marine environments. A multifaceted approach to the study of chemical fossils should greatly improve our knowledge of the evolution and composition of the microbial communities which dominated the biosphere for the greater part of Earth's history.

We thank V. Grey for technical assistance, Z. Roksandic for isotope analyses and J. Bauld and M. R. Walter for helpful comments. Initial work was carried out by T.G.P. while employed by the Geological Survey of Canada in Calgary. This paper is published with the permission of the Director, Bureau of Mineral Resources, Geology and Geophysics.

Received 13 August; accepted 5 December 1985.

1. Mackenzie, A. S. in *Advances in Petroleum Geochemistry* Vol. 1 (eds Brooks, J. & Welte, D. H.) 115–214 (Academic, London, 1984).
2. Stahl, W. J. *Geochim. cosmochim. Acta* **42**, 1573–1577 (1978).
3. Sofer, Z. *Bull. Am. Ass. petrol. Geol.* **68**, 31–49 (1981).
4. Powell, T. G., Macqueen, R. W., Barker, J. F. & Bree, D. G. *Bull. Can. petrol. Geol.* **32**, 289–312 (1984).
5. Powell, T. G. in *Studies in Geology* No. 18 (ed. Palacas, J. G.) 45–62 (American Association Petroleum Geologists, Tulsa, 1984).
6. Ostroukhov, S. B., Arefev, O. A., Makushina, V. M., Zabrodina, M. N. & Petrov, A. A. *Neftekhimiya* **22**, 723–728 (1982).
7. Schaeffe, J., Ludwig, B., Albrecht, P. & Ourisson, G. *Tetrahedron. Lett.* **41**, 3673–3676 (1977).
8. Liaaen-Jensen, S. in *Marine Natural Products* (eds Faulkner, D. J. & Fenical, W. H.) 1–73 (Academic, New York, 1978).
9. Imhoff, J. F. & Truper, H. G. *Microbiol. Ecol.* **3**, 1–9 (1976).
10. Imhoff, J. F. & Truper, H. G. *Zbl. Bakt. I. Abt. Orig.* **C1**, 61–69 (1980).
11. Liaaen-Jensen, S. in *Photosynthetic Bacteria* (eds Clayton, R. K. & Sistrom, W. R.) 233–248 (Plenum, New York, 1978).
12. Liaaen-Jensen, S., Hegge, E. & Jackman, L. M. *Acta chem. Scand.* **18**, 1703–1718 (1964).
13. Hall, P. B. & Douglas, A. G. in *Advances in Organic Geochemistry 1981* (eds Bjoroy, M. et al.) 576–587 (Wiley, Chichester, 1983).
14. Degens, E. T., Behrendt, M., Gothardt, G. & Reppmann, E. *Deep-Sea Res.* **15**, 11–20 (1968).
15. Truper, H. G. & Genovesi, S. *Limnol. Oceanogr.* **13**, 225–232 (1968).
16. Abella, C., Monesinos, E. & Guerrero, F. *Dev. Hydrobiol.* **3**, 173–181 (1980).
17. Fuller, R. C. in *The Photosynthetic Bacteria* (eds Clayton, R. K. & Sistrom, W. R.) 691–705 (Plenum, New York, 1978).
18. Deines, P. in *Handbook of Environmental Isotope Geochemistry* Vol. 1 (eds Fritz, P. & Fontes, J.) 329–406 (Elsevier, Amsterdam, 1980).
19. Schidlowski, M., in *Cosmochemistry and the Origin of Life* (ed. Ponnampetuma, C.) 277–299 (Dordrecht, Reidel, 1983).
20. Schidlowski, M., Matzigkeit, U. & Krumbein, W. E. *Naturwissenschaften* **71**, 303–308 (1984).
21. O'Leary, M. H. *Phytochemistry* **20**, 553–567 (1981).

The Pliensbachian and Tithonian extinction events

A. Hallam

Department of Geological Sciences, University of Birmingham,
PO Box 363, Birmingham B15 2TT, UK

The claim by Raup and Sepkoski¹, that an approximately 26-Myr mass extinction periodicity can be recognized on a global scale at family level over the past 250 Myr, has excited much attention and provoked speculation about an ultimate extraterrestrial cause^{2–10}. While the purported periodicity can be challenged on grounds of both the inaccuracy of the geological timescale¹¹ and the method of extinction analysis¹², there can be no doubting the reality of at least some of the 12 proposed extinction events, such as those at or near the end of the Permian, Triassic and Cretaceous. Others are less widely acknowledged. Further analysis by Raup and Sepkoski¹³ confirms the importance of two of these events within the Jurassic, at the end of the Pliensbachian and Tithonian stages. I demonstrate here that marine invertebrate mass extinctions at approximately these times occurred on a regional, not a global scale. They can be related to severe reductions in habitat area provoked either by regression of epicontinental sea or by widespread anoxia.

As interest in mass extinctions has increased over the past few years, a dearth of detailed documentation has become

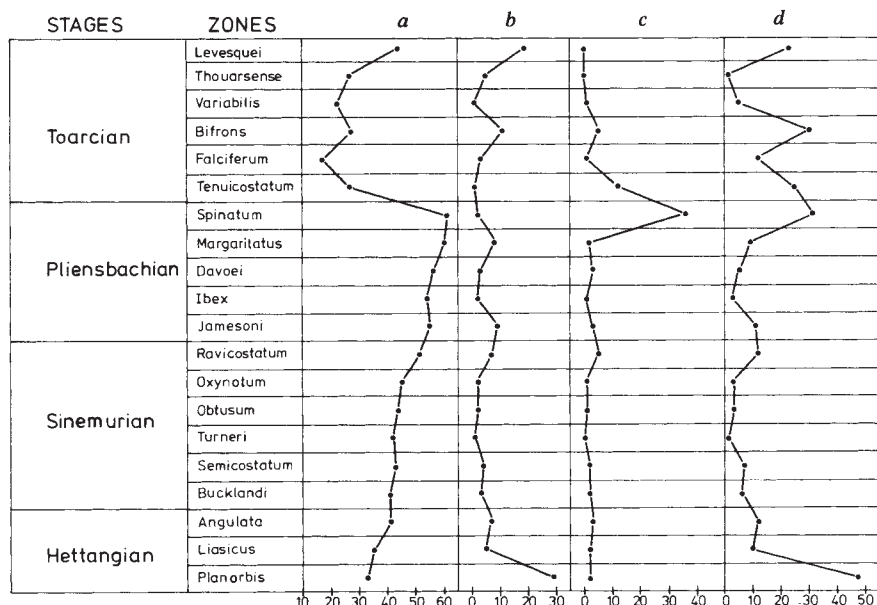


Fig. 1 Patterns of stratigraphic change in the European Lower Jurassic bivalve species, based on method of analysis outlined in ref. 14, and unpublished data on zonal distributions. *a*, Diversity; *b*, time of first appearance; *c*, time of last appearance; *d*, % turnover rate. The drastic reduction in diversity from the late Pliensbachian to the early Toarcian portrayed in *a* is the result of the striking extinction event portrayed in *c*. The turnover rate (*d*) is determined by averaging the number of first and last appearances and dividing by the total number of species.

apparent for all but a few extinction events. One important unresolved question is whether apparently catastrophic events like those at the end of the Permian and Cretaceous were different qualitatively as well as quantitatively from the more numerous smaller-scale events such as those discussed here, which have been studied far less. Finer-grained analyses than those of Raup and Sepkoski are required, down to species level where possible, at the maximum stratigraphic precision available. There should also be thorough documentation of changes in the stratal sequence in different regions, in the hope of detecting correlations with biotic change that may lead to the inference of causal relationships. The present study focuses on the molluscs, especially bivalves, which make up the bulk of the macroinvertebrate fauna in two regions widely separated geographically and with markedly different geological histories: Western Europe and the southern Andes.

Pliensbachian. Because of their high rate of turnover through time, ammonites are in some respects less useful in recording significant extinction 'signals' in the Jurassic than are other common groups that normally exhibit little significant change from stage to stage, such as the bivalves, brachiopods, ostracods and foraminifera. Among the bivalves, the extinction event at or close to the end of the Pliensbachian was by far the most important of the whole Jurassic, with 84% of species becoming extinct in Western Europe, although no such striking change is discernible at the generic level¹⁴. The rhynchonellid brachiopods in Great Britain show an equally striking change, with none of the 35 Pliensbachian species recorded as crossing the Pliensbachian–Toarcian boundary, and 18 of these disappearing at the end of the stage¹⁵. The ostracod faunas in Western Europe also reveal a profound extinction event, one of the most important in the whole post-Palaeozoic history of the group^{16,17}. Moreover, a pronounced turnover of foraminiferal species at approximately the level of the stage boundary is recorded from the Paris Basin¹⁸, although the change is apparently less marked in Great Britain¹⁹.

Stratigraphic analysis within the Lias of Europe can be conducted for most groups at the level of zone rather than stage. By using a similar method of analysis to that presented in ref. 14, it is possible to demonstrate a dramatic increase in bivalve extinction rate at the end of the Pliensbachian (Spinatum Zone) and beginning of the Toarcian (Tenuicostatum Zone; Fig. 1). Breakdown of the bivalve species into ecological groups fails to reveal any selectivity in the extinctions. The most significant facies change corresponding to the extinction event is not that indicating widespread shallowing and regression at the end of

the Pliensbachian, so much as a very extensive and distinctive horizon of laminated organic-rich shale in the lower part of the Falciferum Zone (Lower Toarcian). This horizon is very extensive in Western Europe, being known in different countries under various names, such as Jet Rock, schistes cartons and Posidonienschiefer, and signifies an episode of widespread anoxia in the bottom waters^{20–22}. Many invertebrate species disappear at the often knife-sharp boundary at the base of the Falciferum Zone. When normal shales return in the later Toarcian (Bifrons Zone), new faunas start to enter which bear little close relationship to the older faunas.

All the groups mentioned so far are benthic, but the nektonic or nektobenthic ammonites and belemnites (P. Doyle, personal communication) also exhibit pronounced turnover. The most profound change among the ammonites is the extinction of the boreal amaltheid family at the end of the Pliensbachian, and the emergence at the beginning of the Toarcian of dactylioceratids and hildoceratids of Tethyan derivation. Further important changes among the latter groups took place across the Tenuicostatum–Falciferum Zone boundary. The extinction of the amaltheids, and of those bivalve and other benthic species which disappeared in the Spinatum Zone, appears to be related to marine shallowing caused by regression²⁰, which seems to have been significant on a regional rather than a global scale²³. However, there is probably a slight bias towards the end-Pliensbachian rather than the beginning-Toarcian extinctions because of the greater volume of strata in the Spinatum than the Tenuicostatum Zone, so that further collecting may reveal a few more species surviving the stage boundary.

It is significant that the only marine organisms unaffected by the anoxic event in Europe are planktonic. Of 11 coccolith species recorded from the Pliensbachian, 9 range through the whole Toarcian as well, and the other 2 disappeared well before the end of the Pliensbachian²⁴. Similarly, the 2 species of pseudo-planktonic crinoids survived the anoxic event whereas the more numerous benthic species did not²⁵. These facts appear to indicate that the environmental deterioration caused by the anoxia did not extend into surface waters.

The early Toarcian black shale horizon in Western Europe relates to a global rise of sea level, but although black shales of this age also occur in North America and Japan, the evidence does not support a global anoxic event at this time²⁶. In particular, there is no important black shale horizon in South America^{27–29}, nor is there evidence in this region of a correlative extinction event among bivalves or other groups³⁰. Thus, the abundant Sinemurian–Pliensbachian pectinid genus *Weyla* per-

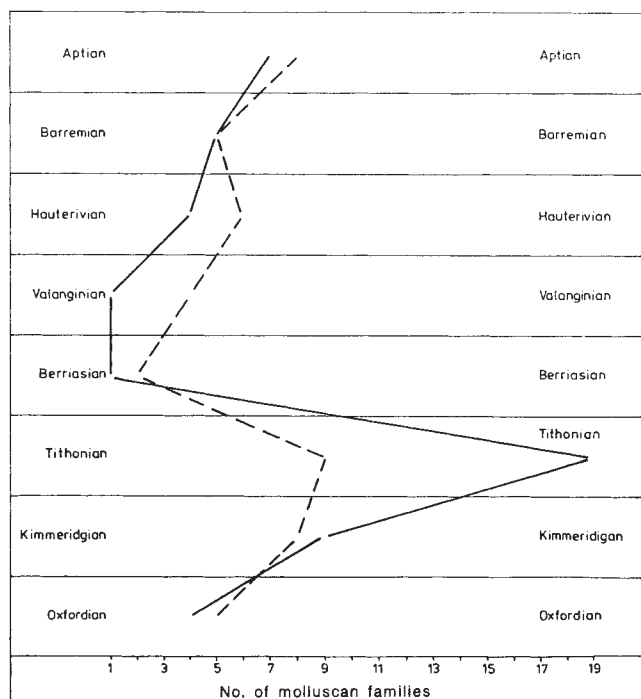


Fig. 2 Extinction (—) and origination (---) rates of late Jurassic and early Cretaceous molluscan families in the world, expressed as times of last and first appearance, respectively. (Based on data in ref. 31.)

sisted until the late Toarcian, as did other endemic forms such as *Plicatostylus* whose inferred ecology suggests relatively high vulnerability to environmental disturbance. Further evidence suggests that Andean South America, and indeed the Pacific margins in general, was little affected by the early Toarcian extinction event. The bivalves that entered the European stratal succession from late Toarcian times onwards can for the most part be matched at genus and even species level with forms that had been confined to the Pacific margins until the late Pliensbachian³⁰. This suggests trans-global migration of species to occupy niches vacated as a consequence of the extinction event. A similar process of expansion of habitat area following niche vacation by extinction can be inferred for the dactyloceratids and hildoceratids that replaced the amaltheids in Western Europe at the beginning of the Toarcian.

Tithonian. If the claim of a significant extinction peak in the Tithonian is to be sustained, it should be clearly expressed in the dominant marine macroinvertebrate group, the Mollusca. Figure 2, based on Sepkoski's compilation³¹ of Phanerozoic families, shows that this is indeed the case. The Tithonian also marks an origination peak, which is consistent with the idea that this stage lasted longer than the immediately subjacent and suprajacent stages; however, the origination peak is much lower than the extinction peak, which therefore emerges as a clear signal. The broad correlation of extinction and origination rates through the eight stages of Fig. 2 reinforces the criticism¹¹ that the Harland timescale, which assumes equal stage durations within the Mesozoic, is not the most reliable for Raup and Sepkoski's cyclicity analysis^{1,13}, which assumes that family extinctions were concentrated at the end of each stage. Only for the ammonites is there sufficiently precise stratigraphic information to test this assumption.

Sepkoski³¹ records that 7 out of a total of 11 ammonite families became extinct in the Tithonian. However, the literature recording more precise stratigraphic information^{32,33} reveals that only 3 of the 7, the Aspidoceratidae, Ataxioceratidae and Dorsoplanitidae, persisted to the end of the stage, and that among these only a minority of genera persisted. The duration of the Tithonian is poorly known but, whatever modern timescale is

used, 5 Myr seems to be a reasonable minimum estimate³⁴. Thus what may appear superficially to be an important extinction event (7 families out of 11) turns out on closer analysis to be of no great significance compared with other extinction events in ammonite history. Indeed, the turnover of taxa at the Tithonian-Berriasian boundary is less marked than that at the Berriasian-Valanginian boundary both in Europe³⁵ and South America³⁶.

Whereas other groups such as the brachiopods^{37,38}, foraminifera^{19,39} and coccoliths²⁴ show no particularly marked change across the Jurassic-Cretaceous (=Tithonian-Berriasian) boundary, even at species level, there was an important late Tithonian extinction event in Europe among species of the dominant group of benthic invertebrates, the bivalves. Thus in Western Europe, Kelly⁴⁰ finds that by far the highest rate of species turnover between the late Jurassic and early Cretaceous is within the Middle Volgian, which is equivalent to the Upper Tithonian⁴¹, while in the southern European Soviet Union (Crimea and Caucasus) the late Tithonian witnessed a striking bivalve extinction event⁴². This event correlates broadly with evidence of a major regression of shallow epicontinental seas which can be traced all the way across Europe from Great Britain to the Soviet Union⁴³⁻⁴⁵. The relative vulnerability of bivalves to extinction, compared with other groups, can probably be accounted for by the restriction of most species to shallow neritic habitats.

In the Andes of central Chile and Argentina, palaeogeographic relationships were very different from those of Europe, with a thick unit of shallow marine Tithonian and Neocomian sediments sandwiched between non-marine units, testifying to a widespread Tithonian transgression and Barremian-Aptian regression; there is no notable facies change or hiatus across the Jurassic-Cretaceous boundary^{29,36,46}. Of 38 species of bivalve distinguished in the Tithonian to Hauterivian sequence of the Chilean High Cordillera, nearly all occur in the Tithonian and none disappears at or near the Tithonian-Berriasian boundary⁴⁶. In fact, the bivalve fauna shows no change up the sequence that cannot be ascribed to facies differences. Species of such characteristic 'Cretaceous' genera as *Exogyra*, *Pterotrignia*, *Steinmanella*, *Eriphyla*, *Ptychomya*, *Isocardia* and *Panopaea* are already present in the Tithonian, a situation quite unlike that in Europe. Thus, the bivalve extinction event cannot be recognized in South America, which also lacks evidence of a corresponding regression at the appropriate time.

In conclusion, although the 'Pliensbachian' (more accurately late Pliensbachian-early Toarcian) and Tithonian extinction events claimed by Raup and Sepkoski can be recognized at species level for at least some marine groups, they appear to be phenomena of regional, not global extent. The facies correlations strongly suggest that the primary causal factor was severe habitat restriction or deterioration due to either regression or the development of widespread anoxia of bottom waters, both phenomena being bound up with sea-level changes on at least a regional scale.

I thank L. Biró-Bagoczky, E. Perez and V. Covacevich for help in fieldwork in Chile and for making available fossil collections in their care. Financial support is acknowledged from the Royal Society and the Comisión Nacional de Investigación Científica y Tecnológica.

Received 15 August 1985; accepted 15 January 1986.

1. Raup, D. M. & Sepkoski, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801-805 (1984).
2. Rampino, M. A. & Stothers, R. B. *Nature* **308**, 709-712 (1984).
3. Schwartz, R. D. & James, P. B. *Nature* **308**, 712-713 (1984).
4. Whitmire, D. P. & Jackson, A. A. *Nature* **308**, 713-715 (1984).
5. Davis, M., Hut, P. & Muller, R. A. *Nature* **308**, 715-717 (1984).
6. Alvarez, W. & Muller, R. A. *Nature* **308**, 718-720 (1984).
7. Clube, S. V. M. & Napier, W. M. *Nature* **311**, 635-636 (1984).
8. Hut, P. *Nature* **311**, 638-641 (1984).
9. Rampino, M. A. & Stothers, R. B. *Science* **226**, 1427-1431 (1984).
10. Whitmire, D. P. & Matese, J. J. *Nature* **313**, 36-38 (1985).
11. Hallam, A. *Nature* **308**, 686-687 (1984).
12. Hoffman, A. *Nature* **315**, 659-662 (1985).
13. Sepkoski, J. J. & Raup, D. M. in *Dynamics of Extinction* (ed. Elliott, D.) (Wiley, New York, in the press).

14. Hallam, A. *Lethaia* 9, 245-259 (1976).
15. Ager, D. V. *A Monograph of the British Liassic Rhynchonellidae*, 1-172 (Palaeontological Society, London, 1958-1967).
16. Plumhoff, F. *Senckenberg. leth.* 48, 549-577 (1967).
17. Lord, A. R. *Palaeontology* 17, 599-622 (1974); in *Aspects of Micropalaeontology*, (Allen & Unwin, London, 1982).
18. Bizon, G. *Mém. Bur. Rech. géol. miner.* 4, 109-114 (1961).
19. Jenkins, D. G. & Murray, J. W. (eds) *Stratigraphic Atlas of Fossil Foraminifera* (Horwood, Chichester, 1981).
20. Hallam, A. *Phil. Trans. R. Soc. B252*, 292-445 (1967).
21. Küspert, W. in *Cyclic and Event Stratification* (eds Einsele, G. & Silacher, A.) 482-501 (Springer, Berlin, 1982).
22. Clayton, C. & Jenkins, H. C. *Sedimentology* 32 (in the press).
23. Hallam, A. *Rev. Earth planet. Sci.* 12, 205-243 (1984).
24. Hamilton, G. B. in *A Stratigraphic Index of Calcareous Nannofossils* (ed. Lord, A. R.) 17-39 (Horwood, Chichester, 1982).
25. Simms, M. J. *Palaeontology* 29 (in the press).
26. Hallam, A. *J. geol. Soc.* 138, 735-743 (1981).
27. von Hillebrandt, A. *Münster Forsch. Geol. Paläont.* 31/32, 167-199 (1973).
28. Perez, d'A. E. *Boln Serv. nac. geol. miner.* 37, 1-149 (1982).
29. Malumian, N., Nullo, F. E. & Ramos, V. A. in *The Phanerozoic Geology of the World Vol. 2B* (eds Moullade, M. & Nairn, A. E. M.) 265-304 (Elsevier, Amsterdam, 1983).
30. Hallam, A. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* 43, 181-193 (1983).
31. Sepkoski, J. J. *Milwaukee Publ. Mus. Contr. Geol. Biol.* 51, 1-125 (1983).
32. Moore, R. C. (ed.) *Treatise on Invertebrate Palaeontology* Vol. 4 (University of Kansas Press, 1957).
33. Donovan, D. T., Callomon, J. H. & Howarth, M. K. in *The Ammonoidea* (eds House, M. R. & Senior, J. R.) 101-156 (Academic, London, 1981).
34. Hallam, A. *Proc. 27th int. geol. Congr., Moscow*, Vol 1, 189-212 (VNU Science Press, Utrecht, 1984).
35. Allemann, F., Grün, W. & Wiedmann, J. *Mém. Bur. Rech. géol. miner.* 86, 14-22 (1975).
36. Leanza, H. A. *Neues Jb. Geol. Paläont. Abh.* 161, 62-92 (1981).
37. Ager, D. V. *Mém. Bur. Rech. géol. miner.* 86, 150-162 (1975).
38. Dagis, A. S. in *The Jurassic-Cretaceous Boundary and the Berriasian Stage in the Boreal Realm* (ed. Saks, V. N.) 244-248 (Acad. Sci. USSR; Israel program for scientific translations, Jerusalem, 1975).
39. Basov, V. A., Balyannikova, S. P., Gorbachik, T. N. & Kuznetsova, K. I. *Mém. Bur. Rech. géol. miner.* 86, 163-8 (1975).
40. Kelly, S. R. A. thesis, Univ. London (1977).
41. Zeiss, A. *Zitteliana* 10, 427-438 (1983).
42. Zakharov, V. A. & Yanine, B. T. *Mém. Bur. Rech. géol. miner.* 86, 221-228 (1975).
43. Hallam, A. *Palaeogeogr. Palaeoclimatol., Palaeoecol.* 23, 1-32 (1978).
44. Ager, D. V. *The Geology of Europe* (McGraw-Hill, London, 1980).
45. Oschmann, W. *Münchn. geowiss. Abh.* A2, 1-119 (1985).
46. Hallam, A., Biró-Bagocsky, L. & Perez, E. *Geol. Mag.* 123 (in the press).

Trampling as a cause of bone surface damage and pseudo-cutmarks

Anna K. Behrensmeyer*, Kathleen D. Gordon†
& Glenn T. Yanagi*

* Department of Paleobiology, NHB-E207 MRC 121,
Smithsonian Institution, Washington, DC 20560, USA

† Department of Anthropology, NHB-363 MRC 112,
Smithsonian Institution, Washington, DC 20650, USA

There have been many recent observations of trampling and its effect on bone surfaces¹⁻⁸ as well as some experimental investigation of the process⁹⁻¹⁴. Although there is known to be a relationship between trampling and scratches on bones, there has been no detailed microscopic comparison with marks made by stone tools. As distinguishing cutmarks from other types of surface features is important in interpreting early hominid behaviour^{9,14-18} and the entry of humans into the New World^{7,19-20}, we have examined possible mimics caused by trampling in an attempt to define diagnostic criteria for cutmarks. We find that microscopic features alone are not sufficient evidence to distinguish human-generated cutmarks from the results of trampling. We suggest different lines of evidence which together may achieve this goal.

We used naturally cleaned bones from domestic horse and cow to test the effects of trampling versus slicing with a stone tool. The bones were relatively fresh (Stage 0-1 weathering³) and were free of periosteum and other tissue. A rib and a metapodial bone were subjected to 3 min of trampling by human feet wearing soft-soled shoes in the damp sand and gravel of a natural stream. The substrate was coarse, sub-angular sand (Mode 1.0-2.0 mm) with some larger grains 5 mm in diameter (Fig. 1a).

As controls, surfaces were replicated using previously described techniques²¹ before the experiment, after cleaning with acetone, water and a soft brush. Specific areas were outlined with cutmarks using a chalcedony flake tool, leaving marks with fine internal striae that are typical of slicing with stone tools²²⁻²⁴ (Figs 2a, c, g; 3a, c). The same areas were cleaned and replicated after trampling. Bone surfaces with cutmarks were also replicated before and after very gentle cleaning with a soft brush and water to determine whether some of the diagnostic microscopic features could be formed on easily removed bone debris left in and beside cutmarks by the slicing process. A single 4 mm grain was mounted in epoxy resin on a wooden block and also used to produce individual scratches on bone surfaces (Fig. 2c, d).

The bones acquired many obvious scratches during trampling (Fig. 2d-f, h) which were most pronounced on the upturned surfaces where grains were pushed across the bone surface by the pressure of the moving foot. Repeated experiments showed

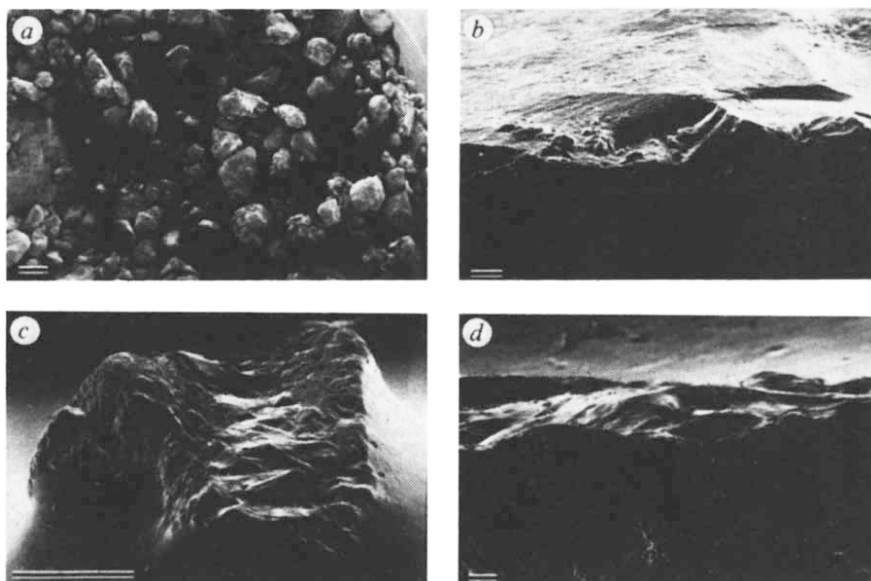


Fig. 1 a, Sand used as substrate during a 3-min trampling experiment; b, edge of chalcedony flake used to make slicing marks on bones; c, single sand grain used to mark bones, mounted in epoxy; d, functional edge of sand grain shown in c for comparison with tool edge (b).

Scale bars: a, c, 1 mm; b, d, 100 μ m.

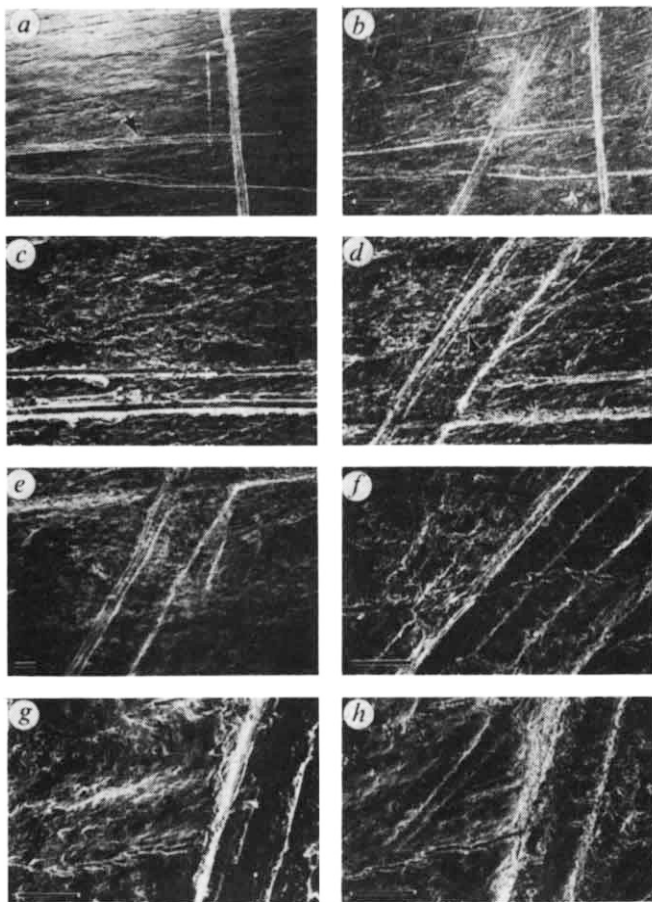


Fig. 2 *a*, Slicing marks made with chalcedony tool on clean concave surface of rib; *b*, same area as *a* after 3 min trampling in sand shown in Fig. 1*a*; *c*, enlarged view of area near the centre of horizontal slicing marks in *a* (arrow) before trampling; *d*, same area as *c*, after trampling (note crack marked by arrow); *e*, set of sub-parallel grooves caused by trampling—original horizontal cutmark appears as rounded groove across the top of the photo; *f*, enlarged view of trample marks in area indicated by arrow in *d* (note crack). Depth of marks ~ 0.1 mm; *g*, enlarged view of pre-trampling slicing mark and bone surface; *h*, same area as *g* after trampling, showing loss of internal detail in the slicing mark and new parallel, V-shaped marks caused by trampling. Depth of slicing mark ~ 0.1 mm. Scale bars: *a*, *b*, 1 mm; *c*–*h*, 100 μ m.

that the metapodials tend to turn around their long axis, thus acquiring scratches oriented normal to this axis. Ribs flip over when stepped on near the ends but usually remain stable, concave side up, when stepped on near the centre. Scratches are concentrated towards the centre of curvature of the rib as a whole, oriented perpendicular or oblique to the long axis.

Examination in the scanning electron microscope (SEM) showed that the scratches occur singly or as sets of parallel marks, some with internal grooving revealed at high magnification (Fig. 2*f*). The marks have both V-shaped and rounded basal cross-sections, and the outer edges are generally rounded. Variability in the sharpness of the trample marks is partly caused by subsequent secondary abrasion during the same episode of trampling. SEM examination also revealed that trampling had significantly altered the features of the pre-existing cutmarks (Fig. 2*b*, *d*, *e*, *h*), eliminating the original sharpness and internal striae.

A double cutmark (Fig. 2*c*) was transformed into two grooves with a rounded, rough cross-section (Fig. 2*d*) and is virtually indistinguishable from new marks of the same size caused by trampling (Fig. 2*d*). The latter are more angular in cross-section,

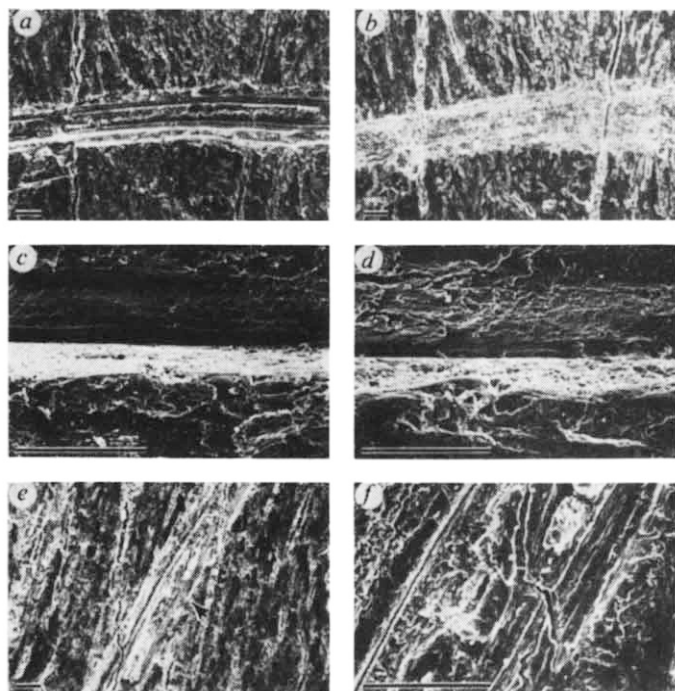


Fig. 3 *a*, Slicing mark from chalcedony flake, before washing; *b*, same area as in *a* after gentle washing with a soft brush and water; *c*, enlargement of V-shaped slicing mark parallel to bone fibre direction before washing; *d*, same area as *c* after gentle washing showing loss of 'smeared' bone debris which bears the fine striae in *c*; *e*, grooves with internal striations made with a single sediment grain on a rib surface (no washing treatment of specimen). Depth of scratch, 0.4 mm; *f*, enlarged view of area in centre of *e* (arrow). Scale bars, 100 μ m.

smoother and better defined than the cutmark. Another cutmark (Fig. 2*g*, *h*) showed loss of its original sharp edge and internal grooving. A pair of clearly defined, V-shaped, sub-parallel scratches acquired during trampling occur next to the abraded cutmark (Fig. 3*h*).

Gentle washing of cutmarks not subjected to trampling showed that most of the internal striations that are clearly visible before washing (Fig. 3*a*, *c*) were removed by this process, leaving rough-surfaced grooves in the bone which resemble cutmarks after trample abrasion (Fig. 2*b*, *d*). This implies that one of the diagnostic features of cutmarks on recent bones^{22,23} can be formed in softened bone debris rather than incised into the bone itself. Striae formed in such debris could easily be removed by natural water action, trampling or other preburial taphonomic processes. Similar loss of striae may also occur during boiling to remove periosteum from fresh bone (Figs 5*a*, *b* and 9*a*, *b* of ref. 23). Strongly irregular tool edges and/or a particularly forceful slicing action may be required to incise striations into the bone itself.

Scratching of bone surfaces with a single-sediment grain produced grooves with irregular cross-sections and internal striations (Fig. 3*e*, *f*) similar to slicing marks made with the stone tool before washing.

The grooves and scratches acquired during trampling were caused by individual and clustered grains of sand with sharp, irregular edges sliding over the bone surface under pressure. The single-grain experiment further demonstrated that sediment grain scratching can create microscopic features which closely resemble the effects of slicing with a stone tool. In the context of each individual groove, sediment scratching can be similar to tool use in that a sharp, irregular edge is brought into forceful contact with the bone surface, as also noted by Oliver⁹.

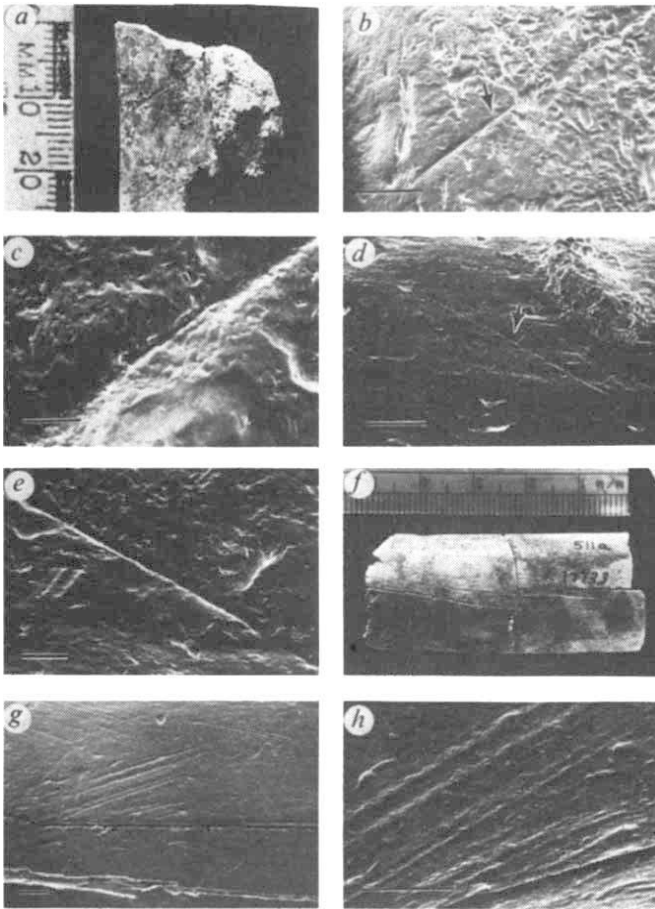


Fig. 4 *a*, Fossil rib section with two isolated parallel, V-shaped marks oblique to the shaft. (Specimen W2-554, locality Y311, 8.8 Myr BP); *b*, enlarged view of lower mark in *a* also showing partial dissolution of bone surface after burial and fossilization; *c*, detail of mark near centre of photograph in *b* (arrow) showing deeper internal groove. (Note similarity to groove in Fig. 3*d*.) *d*, Fossil rib with deep scratch as part of a set of parallel marks oblique to the fibre direction of bone (Specimen W2-507, locality Y311, 8.8 Myr BP. Matrix on upper right includes sand grains comparable in size to the scratches. *e*, Detail of area in centre of photograph in *d* (arrow) showing well-defined, rounded groove; *f*, section of fossil bovid radius shaft showing many clusters of parallel grooves (Specimen W2-511a, locality Y311, 8.8 Myr BP); *g*, enlarged view of area to left of centre in *f*; *h*, detail of scratches just above centre in *g*. Smooth irregular pits are caused by post-fossilization dissolution. Scale bars: *c*, *e*, 100 μ m; *b*, *d*, *g*, *h*, 1 mm.

Experiments by Shipman and Rose²³ and Bromage²⁵ have demonstrated that abrasion by sedimentary particles can significantly alter bone surfaces within minutes to hours. Trampling abrasion and particle flow abrasion (that is, caused by contact with a moving stream of particles in water or air) differ in that trampling can exert more instantaneous force in small areas and can be more variable in contact events with bones. Our experiments show that abrasion under pressures exerted by trampling can alter pre-existing features more rapidly and completely than other types of sedimentary abrasion tested so far.

In a pre-hominid Miocene vertebrate locality in Pakistan, primary scratches and grooves occur on 40–50% of the well-preserved bone surfaces²⁶. Scratches typically occur as sets of shallow, parallel marks oriented oblique or perpendicular to the long axis of rib or limb shafts (Fig. 4*d–h*) or as isolated grooves (Fig. 4*a–c*). The overall pattern of the scratches and various aspects of their internal morphology resemble marks produced during the trampling experiments on recent bones.

Fiorillo^{11–13} has also documented the presence of sets of parallel scratches and grooves on bones from a North American Miocene excavation and has shown how similar scratches were produced on recent bones exposed to natural trampling. Oliver^{8,9} has attributed grooves and other marks on Holocene bones to trampling on a cave substrate with chert-bearing limestone blocks, and Haynes and Stanford⁷ have interpreted trampling as a cause of linear marks on bones from North American Pleistocene localities.

Cutmarks have been reported on a number of African Plio-Pleistocene fossil assemblages^{15–17,27} and on some European and North American assemblages^{7,18–20,24}. Published photographs show many single and multiple grooves, some with internal striations and shoulder marks. At comparable magnifications, marks shown by Bunn (Figs 1*c*, 2*a* and *c* of ref. 17), Shipman and Rose (Figs 1 and 2 of ref. 24) and Shipman *et al.* (Fig. 1*b*, *d* of ref. 18) are similar to trample scratches reported in this paper. Other marks, such as shown in Potts and Shipman (Fig. 2 ref. 16), Bunn (Figs 1*a*, *b*; 2*a*, *b* ref. 17) and Potts (Fig. 4.6 of ref. 27) differ from any of the trample marks produced during our experiments.

Correct identification of cutmarks produced by humans rests on the logic that if an unknown mark on a bone resembles one made by a tool, and if no other process makes a similar mark, then the unknown mark was also made by a tool²³. In the experiment reported here, marks made by sand grains during a brief period of trampling are similar, at magnifications up to 400 $\times$, to marks made by a stone tool. Therefore, cutmark mimics can occur on bones subjected to trampling and original true cutmarks can be obscured by the same process.

There are contextual criteria which can be used to argue against trampling as a cause of marks on bones. The absence of sand and gravel from the environment of burial is one, although this cannot eliminate trample modification elsewhere before final burial or the impact of a sand grain embedded in a hoof (R. Haynes, personal communication). A burial context that indicates no bioturbation (that is, trampling), movement or abrasion of bones is another, and this might occur due to rapid autochthonous burial (for example, ash fall, fine-grained flood deposition over stabilized bones or partial carcasses).

Patterns of individual and clustered marks provide another line of evidence. Experimental trampling marks and scratches attributed to trampling on fossil bones reveal a strong tendency towards multiple, parallel, shallow striations oriented oblique or transverse to bone axes^{9–13,26}. A few of the trample marks in our experiments and on the Miocene fossil bones (5–10%) are deeper, isolated grooves that more closely resemble slicing marks. Cutmarks identified on fossil bones from Olduvai Gorge occur as single deep grooves with internal striae and as clusters of similar, subparallel grooves^{15,16,27}. Sub-parallel marks that microscopically appear to be made by repeated interaction with the same sharp edge (for example, Fig. 4.6*a* ref. 27) are strong evidence for human tool use as it is very unlikely that trampling will produce this pattern.

Placement of cutmarks has been suggested as indicative of hominid behaviour when the marks are focused on areas of ligament attachment^{7,14,15,27–29} or occur in recessed areas unlikely to be exposed to other processes³⁰. The hominid-scavenging hypothesis of Shipman²⁸ is based in part on the occurrence of cutmarks on non-meat bearing areas of bones such as metapodial shafts. Trampling preferentially affects flat, rounded and/or convex areas of bones such as limb shafts²⁶. Although it is unlikely that multiple trample scratches could be mistaken for slicing or scraping marks, isolated trample marks on such areas could be misidentified as cutmarks.

Because the incidence of identified cutmarks, even in large bone assemblages, is very low (for example, on only 1–5% of bones examined in refs 17, 27), misidentification of even a few trample marks as cutmarks could have serious consequences for the interpretation of early hominid behaviour. Microscopic

features of individual marks alone provide insufficient evidence for tool use versus trampling. If such evidence is combined with criteria based on context, pattern of multiple marks and placement on bones, however, it should be possible to distinguish the two processes in at least some cases bearing on early human behaviour.

We thank the Smithsonian Natural History SEM staff and P. Andrews, T. Bromage, G. Haynes, R. Potts and P. Shipman for comments on early drafts of this paper.

Received 5 August; accepted 9 December 1985.

1. Brain, C. K. S. *Af. J. Science* **63**, 97-99 (1967).
2. Gifford, D. P. thesis, Univ. California Berkeley (1977).
3. Behrensmeier, A. K. *Paleobiology* **4**, 150-162 (1978).
4. Hill, A. P. *Paleobiology* **5**, 261-274 (1979).
5. Myers, T. P., Voorhies, M. R. & Corner, R. G. *Am. Antiq.* **45**, 483-490 (1980).
6. Behrensmeier, A. K. & Dechant Boaz, D. E. in *Fossils in the Making* (eds Behrensmeier, A. K. & Hill, A. P.) 72-92 (Univ. Chicago Press, 1980).
7. Haynes, G. & Stanford, D. *Quat. Res.* **22**, 216-230 (1984).
8. Oliver, J. S. *Abstr. First Int. Conf. Bone Modific.* 27 (Center for Study of Early Man, Orino, Maine, 1984).
9. Bunn, H. T. thesis, Univ. California Berkeley (1982).
10. Fiorillo, A. R. *Abstr. First int. Conf. bone Modific.* 11 (Center for Study of Early Man, Orino, Maine, 1984).
11. Fiorillo, A. R. *Curr. Res.* **1**, 47-48 (1984).
12. Fiorillo, A. R. in *Proc. of First Intern. Conf. Bone Modific.* (ed. Bonnichsen, R.) (Center for Study of Early Man, Orino, Maine, in the press).
13. Andrews, P. & Cooke, J. *Man* (in the press).
14. Bunn, H. T. *Nature* **291**, 574-577 (1981).
15. Oliver, J. S. in *Proc. First int. Conf. on Bone Modific.* (ed. Bonnichsen, R.) (Center for Study of Early Man, Orino, Maine, in the press).
16. Potts, R. & Shipman, P. *Nature* **291**, 577-580 (1981).
17. Bunn, H. T. in *Animals and Archaeology: I. Hunters and their Prey* (eds Clutton-Brock, J. & Grigson, C.) 21-30. (Brit. Arch. Rep. Int. Ser. 163, 1983).
18. Shipman, P., Fisher, D. C. & Rose, J. J. *Paleobiology* **10**, 358-365.
19. Bonnichsen, R. *Can. J. Anthropol.* **2**, 137-144 (1982).
20. Rancier, J., Haynes, G. & Stanford, D. *Southwestern Lore* **48**, 1-17 (1982).
21. Gordon, K. D. *Am. J. Phys. Anthropol.* **59**, 195-215 (1982).
22. Shipman, P. *Ann. N. Y. Acad. Sci.* **376**, 357-385 (1981).
23. Shipman, P. & Rose, J. J. *J. Anthropol. Archaeol.* **2**, 57-98 (1983).
24. Shipman, P. & Rose, J. J. *J. Archaeol. Sci.* **10**, 465-474.
25. Bromage, T. G. *Am. J. Phys. Anthropol.* **64**, 161-178 (1984).
26. Behrensmeier, A. K., Gordon, K. D. & Yanagi, G. T. in *Proc. First int. Conf. Bone Modific.* (ed. Bonnichsen, R.) (Center for Study of Early Man, Orino, Maine, in the press).
27. Potts, R. thesis, Harvard Univ. (1982).
28. Shipman, P. in *Animals and Archaeology: I. Hunters and their Prey* (eds Clutton-Brock, J. & Grigson, C.) 31-49 (Brit. Arch. Rep. Int. Ser. 163, 1983).
29. Gifford, D. P., Isaac, G. L. & Nelson, C. M. *Azania* **15**, 57-108.
30. White, T. D. in *Achtste Kroon-Voordracht* 5-33 (Spec. Pub. Nederl. Mus. v. Anthropologie en Praehistorie, 1985).

Asymmetry in the evolution of female mating preferences

Jon Seger

Department of Biology, Princeton University, Princeton, New Jersey, 08544, USA

Robert Trivers

Institute of Marine Sciences, University of California, Santa Cruz, California 95064, USA

Trivers¹ has suggested that where genetic or developmental constraints on the expression of a trait prevent male and female fitnesses from being maximized simultaneously, female mating preferences should evolve to favour males who exhibit variants of the trait that confer relatively low fitness on males but relatively high fitness on females. This asymmetry is expected because alleles that affect mating preferences are expressed only in females, but are genetically correlated with alleles that differentially affect the fitnesses of the two sexes. Here we describe a two-locus population-genetic model that embodies this idea. The model's qualitative behaviour is exactly like that of previous models²⁻¹⁰ for the joint evolution of male traits and female mating preferences: evolution is equally likely to proceed in either direction along (or away from) a line of neutral equilibria that relates given frequencies of the preference alleles to corresponding frequencies of the trait alleles.

But there is a quantitative asymmetry, of the expected kind, in the shape of the line of equilibria. When we extend the model to include migration between partially isolated demes (breeding groups), 'selective diffusion'^{11,12} moves the demes along the line of equilibria in the direction that increases average female fitness while lowering average male fitness.

The model species is a sexual haploid with two unlinked loci and discrete, non-overlapping generations. Alleles at the trait locus (T) determine male and female viabilities according to the following scheme:

	T_1	T_2
Males	α_1	α_2
Females	β_1	β_2

In males, T_1 and T_2 also determine a phenotypic difference that is visible to females. At the beginning of each generation the frequencies of T_1 and T_2 are t_1 and $t_2 = 1 - t_1$. After viability selection, the frequencies of T_1 in adult males and females are

$$t_1^m = \alpha_1 t_1 / (\alpha_1 t_1 + \alpha_2 t_2) \quad (1)$$

and

$$t_1^f = \beta_1 t_1 / (\beta_1 t_1 + \beta_2 t_2) \quad (2)$$

with $t_2^m = 1 - t_1^m$ and $t_2^f = 1 - t_1^f$.

Female mating preferences are determined by alleles at the preference locus (P). Females carrying P_1 prefer to mate with T_1 males, and females carrying P_2 prefer to mate with T_2 males. There are many different behavioural models of female choice^{3-10,13,14}; we used two simple ones that give qualitatively different results in some contexts⁸. Under the fixed-relative-preference model^{6,8}, the probability that a P_1 female mates with a T_1 male is

$$U_{11} = a_1 t_1^m / (a_1 t_1^m + t_2^m) \quad (3)$$

where a_1 (≥ 1) is a parameter that sets the strength of the female preference, which vanishes at $a_1 = 1$. The probability that a P_1 female mates with a T_2 male is $U_{12} = 1 - U_{11}$. The behavioural model is one in which P_1 females encounter T_1 and T_2 males sequentially, at random, accepting them as mates with fixed probabilities whose ratio is $a_1 : 1$. Similarly,

$$U_{22} = a_2 t_2^m / (a_2 t_2^m + t_1^m) \quad (4)$$

is the probability that a P_2 female mates with a T_2 male, and $U_{21} = 1 - U_{22}$ is the probability that she mates with a T_1 male.

Under the better-of-two model of female choice,

$$U_{11} = t_1^m + c_1 t_1^m t_2^m \quad (5)$$

and

$$U_{22} = t_2^m + c_2 t_1^m t_2^m \quad (6)$$

where c_1 and c_2 ($0 \leq c_i \leq 1$) are parameters that set the strength of the mating preference. Here the behavioural model is that of a lek. Each lek is composed of two males, sampled at random from the population after viability selection. If the lek contains one T_1 and one T_2 male, the female chooses her preferred type with probability $\frac{1}{2}(1 + c_i)$. Thus the preference vanishes at $c_i = 0$.

The frequencies of P_1 and P_2 are p_1 and $p_2 = 1 - p_1$, at the beginning of each generation. Although P_1 and P_2 have no effects on male or female viabilities, their frequencies change slightly in each sex during viability selection at the T locus, because of the genetic covariance (phase disequilibrium) that exists between the two loci. This covariance is maintained, even at gene-frequency equilibrium, by the female mating preferences²⁻¹⁰.

Given the assumptions and definitions described above, analysis of the model proceeds in a straightforward way, along

Fig. 1 Lines of stable equilibria for models in which only one sex is affected by T_1 and T_2 . **A**, Only males are affected. This is Kirkpatrick's⁶ model in which $\alpha_1 = \beta_1 = \beta_2 = 1$, $\alpha_2 = 0.6$, $a_1 = 1$ and $a_2 = 3$. **B**, Only females are affected. Here $\alpha_1 = \alpha_2 = \beta_1 = 1$, $\beta_2 = 0.6$, $a_1 = 1$ and $a_2 = 3$. T_1 goes to fixation when the frequency of P_1 is greater than 0.8 ($p_2 < 0.2$), but T_2 goes to fixation only when P_2 is also at fixation ($p_2 = 1$). If a_2 (the preference parameter) were larger than 3, or if β_2 (the fitness of T_2 females) were larger than 0.6, then the upper end of the line of equilibria would intersect the upper boundary ($t_2 = 1$) at $p_2 < 1$. Conversely, if a_2 were smaller than 3 or if β_2 were smaller than 0.6, then the line would intersect the right-hand boundary ($p_2 = 1$) at $t_2 < 1$, and T_2 would have an interior equilibrium even at fixation for P_2 . In all of the models discussed here, the interior equilibria satisfy the equation

$$p_2^f = \frac{t_2(2 - 1/F) - U_{12}}{U_{22} + U_{11} - 1} \quad (8)$$

where p_2^f is the frequency of P_2 in adult females and $F = \beta_1 t_1 + \beta_2 t_2$ is the average female fitness. Equation (8) was used to draw the lines of equilibria shown above and in Fig. 2. If female fitnesses are not affected by T_1 and T_2 , then $F = 1$, and $p_2^f = p_2$. But if the T locus is expressed in females ($\beta_1 < 1$ or $\beta_2 < 1$), then $F < 1$, and p_2^f is slightly different from p_2 . The change in the frequency of P_2 in females arises during viability selection at the T locus, owing to the phase disequilibrium (covariance) between the two loci. Thus, the line of equilibria shown in **A** (Kirkpatrick's model, in which $\beta_1 = \beta_2 = 1$) is exact, but the lines shown in **B** and Fig. 2 are approximate. (If the horizontal axes were relabelled as p_2^f , then all of the lines would be exact.) The error involved in this approximation (or misrepresentation) was estimated by numerical iteration of the recurrence equations in the genotype frequencies, which model the full dynamics of the two-locus system. Even with strong selection and strong female mating preferences, as in **B**, p_2^f differs from p_2 by only a few per cent for intermediate gene frequencies, becoming identical to p_2 at the boundaries. (The exact end points of the lines of equilibria can easily be derived analytically.) With weaker selection and weaker mating preferences, as in Fig. 2, the approximation is excellent everywhere.

the lines described more fully elsewhere^{6,8}. Additional details are given in the figure legends.

The asymmetry mentioned above can easily be demonstrated by comparing models in which the fitnesses ($\alpha_1, \alpha_2, \beta_1, \beta_2$) have been symmetrically rearranged. Consider the case $(1, \alpha_2, 1, 1)$, with $\alpha_2 < 1$. Here females are unaffected by their T -locus genotypes, but T_2 males suffer reduced viability relative to T_1 males. Kirkpatrick's⁶ first model is this fitness scheme with fixed-relative-preference female choice, where $a_1 = 1$ and $a_2 > 1$ (P_1 females mate randomly, while P_2 females prefer the less viable T_2 males). Figure 1A shows the line of stable equilibria illustrated in Kirkpatrick's Fig. 1. If we reverse the pattern of sex-limitation, so that T_2 females suffer reduced viability, then the fitness array is $(1, 1, 1, \beta_2)$ and we obtain the line of stable equilibria shown in Fig. 1B. Although T_2 invades at lower frequencies of P_2 when its deleterious effect is limited to females than it does when limited to males, much higher frequencies of P_2 are required to take it to fixation. Of course, T_2 can have any equilibrium frequency between 0 and 1, under either of these fitness schemes, given an appropriate frequency of P_2 . But under almost any probability distribution of $\hat{p}_2$ centred on $\hat{p}_2 = \frac{1}{2}$, the average equilibrium frequency of T_2 is higher when its

deleterious effect is limited to males than when the same deleterious effect is limited to females.

In a more fully symmetrical model, both T_1 and T_2 would affect the fitnesses of their carriers, and both P_1 and P_2 females would exhibit mating preferences. Consider the fitnesses $(1, \alpha_2, \beta_1, 1)$. Here T_1 males are fitter than T_2 males, as in Kirkpatrick's model. But now T_2 females are fitter than T_1 females. If we set $\alpha_2 = \beta_1$, then the fitness effects are completely symmetrical, and in the absence of any female mating preferences, T_1 and T_2 have a strongly attracting equilibrium¹⁵⁻¹⁹ at $\hat{t}_1 = \hat{t}_2 = \frac{1}{2}$. If we also set $a_1 = a_2 > 1$, then P_1 and P_2 females have equally strong preferences for T_1 and T_2 males, respectively, and the model is fully symmetrical. But the line of stable equilibria is not symmetrical. T_2 is taken to high frequencies more easily than is T_1 (Fig. 2).

The interior equilibria are always stable under the fixed-relative-preference model of female choice^{6,8}, but they can be unstable under the better-of-two model⁸. Figure 3 shows a line of equilibria for the symmetrical fitnesses described above, under symmetrical better-of-two female choice ($c_1 = c_2 = 1$). Again the asymmetry is obvious, and is illustrated dramatically by the dynamical behaviour of a population that begins at the centre of the gene-frequency plane ($t_2 = p_2 = \frac{1}{2}$). The population evolves rapidly to a boundary equilibrium at $t_2 = 1$, $p_2 > \frac{1}{2}$ (Fig. 3).

The causes of the asymmetry are easy to see. P_2 females who choose T_2 males instead of T_1 males thereby increase the survival of their daughters. Of course, they also decrease the survival of their sons, so there would seem to be no net advantage associated with the preference for T_2 . But in the next generation it will be females, not males, who do the choosing, and allele P_2 increases its frequency in females by associating itself nonrandomly with T_2 .

The stable equilibria are points at which a population would sit forever, if affected only by the deterministic forces of natural and sexual selection. But in a world of finite, partially isolated demes, differential migration may give rise to a weak directional force that moves an average deme along the line of equilibria, towards higher frequencies of P_2 and T_2 . Consider a set of n demes, each of which has the same carrying capacity, K . The expected number of adult females in the i th deme is $KF_i/(F_i + M_i)$, where F_i and M_i are the average viabilities of females and males born in that deme. If males do not help females to rear offspring, then the productivity of each deme is proportional to its number of adult females, which is proportional to its frequency of T_2 , under the fitness scheme discussed above $[(1, \alpha_2, \beta_1, 1)$ with $\alpha_2 = \beta_1$]. Let each deme contribute to the migrant pool in proportion to F_i , and let it then receive mK migrants from the pool. It follows that

$$\Delta \bar{t}_2 = \frac{m}{1+m} \left[\frac{(1-\beta_1) \text{var}(t_2)}{\beta_1 + (1-\beta_1)\bar{t}_2} \right] \quad (7)$$

For example, if $m = 0.01$, $\beta_1 = 0.8$, $\bar{t}_2 = 0.5$ and $\text{var}(t_2) = 0.01$, then $\Delta \bar{t}_2 = 2.2 \times 10^{-5}$. The overall frequency of T_2 increases because demes with higher than average frequencies of T_2 make larger than average contributions to the migrant pool, which

Fig. 2 Line of stable equilibria for a fully symmetrical version of Kirkpatrick's model. Here $\alpha_1 = \beta_2 = 1$, $\alpha_2 = \beta_1 = 0.8$ and $a_1 = a_2 = 1.05$. Note that the strengths of natural selection and female choice are weaker here than in the examples shown in Figs 1 and 3. T_1 and T_2 have an interior equilibrium at fixation for P_1 ($p_2 = 0$), but T_2 is taken to fixation when P_2 is above a frequency of 0.925. Qualitatively similar lines of equilibria exist even if $\alpha_2 \neq \beta_1$ and $a_1 \neq a_2$.

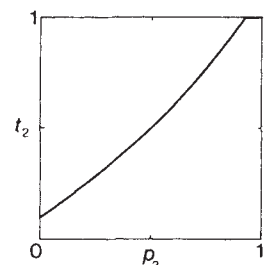
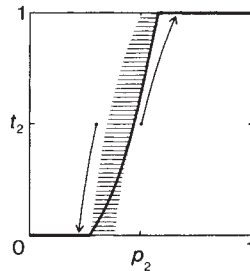


Fig. 3 Line of equilibria for the symmetrical model with better-of-two female choice. Here $\alpha_1 = \beta_2 = 1$, $\alpha_2 = \beta_1 = 0.6$ and $c_1 = c_2 = 1$. From points to the left of the line, populations move down and to the left. From points to the right of the line, they move up and to the right. From points within the hatched region, populations are attracted to stable interior equilibria on the lower region of the line (approximately, the region below $t_2 = 0.55$).



From points outside the hatched region, populations move to the stable boundary equilibria at $t_2 = 0$ or $t_2 = 1$. The illustrated trajectories begin at $p_2 = 0.3$, $t_2 = 0.5$ and at $p_2 = t_2 = 0.5$. Each of these trajectories is less than 100 generations in length. The upper region of the line (between approximately $t_2 = 0.55$ and $t_2 = 1$) is unstable. From points just to the right of the line, populations move to the boundary at $t_2 = 1$, and from points just to the left, they move to stable interior equilibria on the lower region of the line. The line of equilibria, the regions of attraction and the illustrated trajectories were all found numerically, by iteration of the recurrence equations in the genotype frequencies. The approximate line of equilibria (equation (8); see Fig. 1 legend) has the same end points as the line illustrated here, but it passes through the central point ($p_2 = t_2 = \frac{1}{2}$).

therefore has a higher frequency of T_2 than does the population at large.

Suppose that at the beginning of a generation the demes are at various points along an approximately straight segment of a line of stable equilibria. Then after migration, selection and reproduction, the demes will still be on the line, because the migrants that entered each deme will have changed its frequency of P_2 by an amount proportional to the amount by which they changed its frequency of T_2 , the constant of proportionality being the reciprocal of the slope of the line. The changes of gene frequency caused each generation by differential migration are not opposed by selection within demes, because within each deme the new frequency of T_2 is stabilized by a new and appropriate frequency of P_2 . Even where the line of equilibria is moderately curved, migration moves the demes along paths that are almost tangent to the line. Then, depending on the direction of the curvature, there are either slight decreases or slight additional increases in the average frequencies of T_2 and P_2 , as selection pulls the demes in toward the line.

This process of 'selective diffusion'^{11,12} or 'group selection'²⁰⁻²³ depends on the variance of gene frequencies among demes, but the variance is reduced by migration. In group-selection models for the evolution of reproductive altruism or female-biased sex ratios of parental investment, the overall direction of gene-frequency change is reversed when the variance falls below a critical minimum, because within each deme there is selection against the trait that is favoured by differential migration between demes. This problem does not arise in the present model, because within each deme female choice stabilizes the current frequency of T_2 . Thus, $\bar{t}_2$ continues to increase as long as the balance between migration and sampling drift maintains even a small interdemic variance of gene frequencies. Female mating preferences change the frequencies of T_2 only in demes that are away from the line of equilibria. In demes that are on the line they do not directly change gene frequencies, but they can act as a catalyst for changes induced by other causes.

The main implication of these results is that females in polygynous species are expected to prefer mating with males who, other things being equal, show evidence that they would have been reproductively successful had they been females rather than males. We are not the first to argue that female choice should evolve to favour, in part, traits that promote ecological fitness^{1,2,24-34}, but the present model seems to be the

first to show explicitly how female choice could evolve to place special emphasis on traits that promote female fitness^{1,26}.

Many well-known patterns of sexual dimorphism are consistent with an interpretation of the kind implied by this model. For example, rates of senescence are higher in the males of most species than in the females^{35,36}. Unfortunately, this dimorphism has plausible alternative explanations based on male-male competition rather than female choice^{1,35-37}. The same ambiguity seems likely to appear in all comparative tests of the model, as the following example will illustrate.

Male *Anolis* lizards tend to be much larger than conspecific females on islands, but not on the mainland³⁸. Island populations appear to be more strongly food-limited than mainland populations. Individuals belonging to island populations grow relatively slowly, mature at relatively late ages, feed relatively often on small items, and show sharp increases in growth rates when additional food is supplied experimentally³⁸. If the fitnesses of island females depend relatively strongly on the efficient acquisition and assimilation of food, we would expect island females to prefer mating with males who had demonstrated special competence in these abilities. The greater size dimorphism of island populations is consistent with this expectation. But it is also consistent with the argument that female-defence polygyny is a more successful male strategy on islands than it is on the mainland, because island populations live at higher densities (with smaller territory and home-range sizes) than do mainland populations. Direct male-male competition would then be the selective force that favours relatively large male sizes on islands³⁹. The two models are not mutually exclusive, in this example or in general, so considerable ingenuity will be needed to devise tests that clearly distinguish between them.

We thank J. Bull, P. Harvey, R. Lande and M. Kirkpatrick for helpful advice and discussion.

Received 31 October 1985; accepted 6 January 1986.

1. Trivers, R. *Social Evolution* (Benjamin/Cummings, Menlo Park, 1985).
2. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
3. O'Donald, P. *Genetic Models of Sexual Selection* (Cambridge University Press, 1980).
4. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3721-3725 (1981).
5. Lande, R. *Evolution* **36**, 213-223 (1982).
6. Kirkpatrick, M. *Evolution* **36**, 1-12 (1982).
7. Kirkpatrick, M. *Am. Nat.* **125**, 788-810 (1985).
8. Seger, J. *Evolution* **39**, 1185-1193 (1985).
9. Engen, S. & Saether, B.-E. *J. theor. Biol.* **117**, 277-289 (1985).
10. Lande, R. & Arnold, S. J. *J. theor. Biol.* **117**, 651-664 (1985).
11. Wright, S. *Ecology* **26**, 415-419 (1945).
12. Wright, S. *Evolution and the Genetics of Populations* Vol. 3 (University of Chicago Press, 1977).
13. Janetos, A. C. *Behav. Ecol. Sociobiol.* **7**, 107-112 (1980).
14. Parker, G. A. in *Mate Choice* (ed. Bateson, P.) 141-166 (Cambridge University Press, 1983).
15. Owen, A. R. G. *Heredity* **7**, 97-102 (1953).
16. Haldane, J. B. S. *Nature* **193**, 1108 (1962).
17. Li, C. C. *Evolution* **17**, 493-496 (1963).
18. Bodmer, W. F. *Genetics* **51**, 411-424 (1965).
19. Kidwell, J. F. *et al. Genetics* **85**, 171-183 (1977).
20. Wade, M. J. *Q. Rev. Biol.* **53**, 101-114 (1978).
21. Uyenoyama, M. & Feldman, M. W. *Theor. Popul. Biol.* **17**, 380-414 (1980).
22. Wilson, D. S. *A. Rev. Ecol. Syst.* **14**, 159-187 (1983).
23. Wilson, D. S. & Colwell, R. K. *Evolution* **35**, 882-897 (1981).
24. Mayr, E. in *Sexual Selection and the Descent of Man 1871-1971* (ed. Campbell, B.) 87-104 (Aldine, Chicago, 1972).
25. Zahavi, A. *J. theor. Biol.* **53**, 205-214 (1975).
26. Trivers, R. L. *Evolution* **30**, 253-269 (1976).
27. Bell, G. *Evolution* **32**, 872-885 (1978).
28. Eshel, I. *J. theor. Biol.* **70**, 245-250 (1978).
29. Borgia, G. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.) 19-80 (Academic, New York, 1979).
30. Andersson, M. *Biol. J. Linn. Soc.* **17**, 375-393 (1982).
31. Hamilton, W. D. & Zuk, M. *Science* **218**, 384-387 (1982).
32. Dominey, W. J. *J. theor. Biol.* **101**, 495-502 (1983).
33. Thornhill, R. & Alcock, J. *The Evolution of Insect Mating Systems* (Harvard University Press, Cambridge, 1983).
34. Heisler, I. L. *Evolution* **38**, 1283-1295 (1984).
35. Williams, G. C. *Evolution* **11**, 398-411 (1957).
36. Comfort, A. *The Biology of Senescence* 3rd edn (Elsevier, New York, 1979).
37. Rose, M. R. *Am. Zool.* **23**, 15-23 (1983).
38. Andrews, R. M. *Copeia* **1976**, 477-482 (1976).
39. Stamps, J. A. in *Lizard Ecology* (eds Huey, R. B., Pianka, E. R. & Schoener, T. W.) 169-204 (Harvard University Press, Cambridge, 1983).

Selective impairment of learning and blockade of long-term potentiation by an *N*-methyl-D-aspartate receptor antagonist, AP5

R. G. M. Morris*, E. Anderson*, G. S. Lynch† & M. Baudry†

*MRC Cognitive Neuroscience Research Group, University of St. Andrews, St Andrews, Fife KY16 9JU, UK

†Center for the Neurobiology of Learning and Memory, University of California, Irvine, California 92717, USA

Recent work has shown that the hippocampus contains a class of receptors for the excitatory amino acid glutamate that are activated by *N*-methyl-D-aspartate (NMDA)¹ and that exhibit a peculiar dependency on membrane voltage in becoming active only on depolarization^{2,3}. Blockade of these sites with the drug aminophosphonovaleric acid (AP5) does not detectably affect synaptic transmission in the hippocampus, but prevents the induction of hippocampal long-term potentiation (LTP) following brief high-frequency stimulation⁴⁻⁶. We now report that chronic intraventricular infusion of D, L-AP5 causes a selective impairment of place learning, which is highly sensitive to hippocampal damage⁷, without affecting visual discrimination learning, which is not⁸. The L-isomer of AP5 did not produce behavioural effects. AP5 treatment also suppressed LTP *in vivo*. These results suggest that NMDA receptors are involved in spatial learning, and add support to the hypothesis⁹⁻¹¹ that LTP is involved in some, but not all, forms of learning.

On day 0 of the place-learning experiment, male Lister rats ($n=36$) were implanted with minipumps (Alza, 2002) containing D, L-AP5 (40 mM in 0.9% saline, $n=10$; we estimate that the steady-state concentration in cerebrospinal fluid would have risen to 150 μ M), L-AP5 (20 mM, $n=6$) or served as controls made up of saline minipump ($n=10$) and unoperated ($n=10$) subgroups. Using stereotaxic techniques, under Avertin anaesthesia, a cannula (25-gauge) was lowered into the right lateral ventricle (Bregma, -0.9 mm). It was attached to a catheter, secured by dental cement, and connected to the subcutaneously implanted minipump. At the completion of testing, a check made that each pump was still connected to its cannula, and the brains removed for histological verification of the cannula locations.

On day 4, the rats were placed into a large pool of opaque

water (at $26 \pm 1^\circ\text{C}$) and trained to find and escape onto a platform which was hidden at the centre of either the SW or NE quadrants of a large 2.14 m diameter pool¹².

Training was undertaken on days 4-8. We began by giving a total of 15 trials (3 trials per day at a 4-h inter-trial interval (ITI), randomized starting locations, 5 days). The rats swam until they found the platform (maximum swim time 120 s), with 30 s spent on it before removal. Escape latency was measured with a stopwatch. On day 9, a spatial transfer test was performed and further training was given. The purpose of the transfer test was to find out how much had been learned about the location of the escape platform. It was therefore removed from the pool, and the animals given 60 s to swim freely with no opportunity of escape. Their movements were tracked automatically using an image-analysing device¹². Normal rats search near their former location in this kind of test. They were then given eight further training trials in rapid succession (30-s ITI), with the platform back in its original location. In reversal trials, on days 10-13, the rats were trained to learn a new location, in the opposite quadrant from that used in earlier training (only one trial per day).

The control and L-AP5 groups learned to escape relatively rapidly within 15 trials (Fig. 1). The D, L-AP5 group learned more slowly, but their impairment was not striking; considerable

Table 1 Number of trials to the criterion of nine successive correct trials on one day ($P < 0.0005$)

Group	N	Mean	Range
Unoperated	6	60.8*	41-100
Saline	8	80.6	50-120†
D, L-AP5	13	76.8	41-120†
L-AP5	9	67.9	40-100

* The groups did not differ significantly in rate of learning Kruskal-Wallis $H = 7.19$, $P > 0.05$.

† Two saline rats and three D, L-AP5 rats failed to reach criterion before the pumps were exhausted.

'nonspecific' instrumental learning occurs early in training (for example, to swim away from the side walls) and this may have masked an impairment in true place-learning. Moreover, some D, L-AP5 animals occasionally fell off the escape platform on the first 2-3 days of training. However, the outcome of the further training on day 9, after the transfer test, showed that the D, L-AP5 rats stabilized their escape latencies at a much higher level, taking longer, more indirect routes to the platform.

Fig. 1 Mean latency of escape(s) for each of the 10 days of behavioural testing. Days 4-8, mean of the 3 daily trials; day 9, subdividing into 4 blocks of 2 trials; days 10-13, the single reversal trial of each day. The transfer test on day 9 immediately preceded the 8 further training trials. To avoid problems of heterogeneity of variance, the successive phases of the experiment were analysed separately. Training (days 4-8): unweighted means (unequal n) analysis of variance revealed significant effects of group ($F = 7.11$, d.f. 2/33, $P < 0.005$), and days ($F = 101.6$, $P < 0.0001$). Subsequent orthogonal comparisons showed that the escape latency impairment was restricted to the D, L-AP5 group (Control versus L-AP5, $P > 0.10$; D, L-AP5 versus control, $P < 0.01$). Further training (day 9): the groups differed significantly ($F = 7.87$, d.f. 2/33, $P < 0.002$) with the impairment again restricted to the D, L-AP5 group (D, L-AP5 versus control, $P < 0.0005$). Reversal (days 10-13): all groups showed a marked increase in latency on the first trial of reversal, the increases being 75.3 and 53.4 s for the control and L-AP5 groups, but only 34.9 s for the D, L-AP5 group, but this trend did not differ between groups. Over the reversal trials, the groups differed significantly in escape latency ($F = 7.68$, d.f. 2/33, $P < 0.005$) with the impairment in learning to approach the new platform position specific to the D, L-AP5 group. On day 13 group D, L-AP5 escaped more slowly than the L-AP5 and control groups combined ($F = 4.31$, $P < 0.05$). ●, Control; ○, D, L-AP5; □, L-AP5.

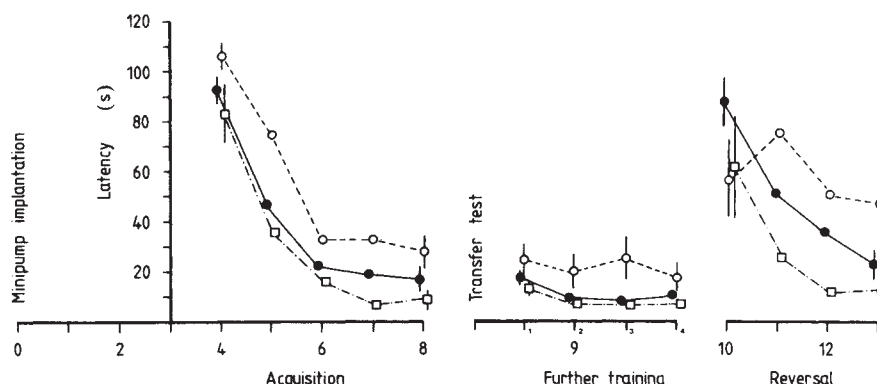
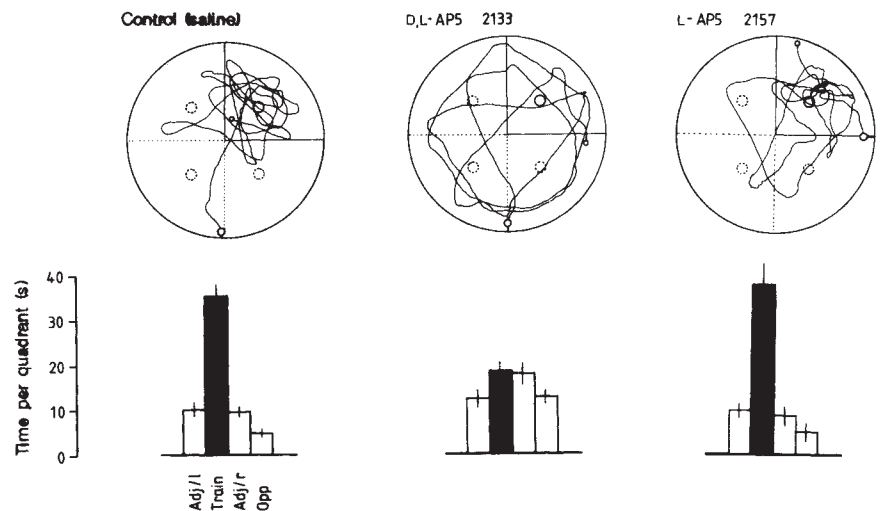


Fig. 2 Swimming path taken by the rat closest to the group mean (calculated according to distribution of time spent per quadrant) in the Control, D, L-AP5 and L-AP5 groups respectively. For each animal, the computerized tracking system calculated the times spent in the four cardinal quadrants of the pool during the 60 s transfer test. The data were organized into time spent in the original training quadrant (Tr), the opposite (Opp) and the two adjacent quadrants (Adj/L, Adj/R). An analysis of the group scores, plotted beneath each representative animal, revealed a highly significant groups $\times$ quadrants interaction effect ($F = 8.62$, d.f. 6/99, $P < 0.0001$). The mean times spent $\pm$ 1 s.e. in the training quadrant were control = 35.4 ± 2.2 s; (59% of the 60-s test where chance = 25%, $n = 20$); D, L-AP5 = 18.2 ± 2.0 s (30%); and L-AP5 = 37.3 ± 4.8 (62%).



L-AP5 and control rats learned to escape with minimum latencies, using direct paths to the platform. During the four reversal trials of days 10–13, the D, L-AP5 group failed to learn the new platform location, while L-AP5 and control animals showed a significant decline in latency across trials.

Analysis of the paths taken during the transfer test on day 9 provided the most convincing demonstration that the D, L-AP5 rats had learned little about the platform location. Figure 2 shows the paths swum by a typical rat of each group, and the mean distribution of time spent in the four quadrants. The D, L-AP5 group showed little spatial bias to the training quadrant.

These data imply that chronic infusion of an NMDA antagonist causes a marked impairment in place learning, similar to that caused by hippocampal lesions⁷. However, we could not distinguish between a primary effect on spatial learning per se, and a secondary effect upon sensorimotor or motivational processes. In a second behavioural study, we used the same apparatus, and thus the same motivation to escape from water, but a learning task which did not require spatial learning.

To assess visual discrimination, two discriminable platforms (grey, and black-and-white stripes) were placed into the pool with their top surfaces 2 cm above the water surface. One was rigid and provided escape; the other was floating and sank under the water whenever the rat attempted to climb on. Black curtains surrounded the perimeter of the pool obscuring extra-maze cues, and the platforms were moved between several locations over successive trials throughout training. The rats' task was to learn to approach the rigid platform irrespective of location. A total of 36 male Lister rats were trained (D, L-AP5, 40 mM, minipumps, $n = 13$; L-AP5, 20 mM, $n = 9$; saline, $n = 8$; unoperated $n = 6$). They were given 10 trials per day (30 s ITI), to a

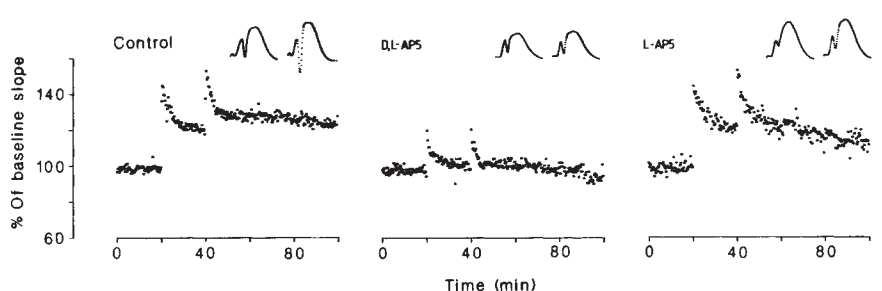
criterion of nine successive correct trials on one day ($P < 0.0005$), starting on day 3 after surgery and continuing until the pumps were exhausted on day 14 (maximum, 120 trials).

The results (Table 1) showed that there was a nonsignificant trend towards faster learning by unoperated control rats but, overall, the groups did not differ. These findings imply that a secondary sensorimotor or motivational impairment is unlikely to be the cause of the place-learning deficit described earlier.

Finally, to determine the effect of AP5 on induction of LTP *in vivo*, we examined LTP using a third set of male Lister rats ($n = 22$), also prepared with intraventricular cannulae and minipumps (D, L-AP5, 40 mM, $n = 9$; L-AP5, 20 mM, $n = 6$; control $n = 7$ made up from saline, $n = 3$, and unoperated $n = 4$). Between 6 and 12 days later, each rat was anaesthetized with urethane (1.5 g per kg, interperitoneally, i.p.) and placed into a Kopf stereotaxic instrument. A bipolar stimulating electrode (110 μ m diameter) was lowered into the perforant path (action potential 7.6 mV, length 4.0 mm); and a bipolar, staggered tip, recording electrode (110 μ m diameter) placed in the dentate gyrus to record both hilus and molecular layer field potentials. After stable potentials had been recorded for at least 30 min, low frequency test pulses (100 μ s, 7 V, 0.05 Hz) were applied for 100 min. Two brief bursts of high-frequency activation (100 μ s, 7 V, 400 Hz, 250 ms) were applied exactly 20 and 40 min into the experiment. We measured the maximum voltage, and the slope of the early rising phase of the extracellular field-potential (1.8–2.5 ms) using linear regression, sampling the waveform at 10 kHz.

D, L-AP5 infusion over 6 to 12 days caused no obvious effects on potentials evoked at low-frequency (pre-tetanus maximum amplitudes were: D, L-AP5 = 7.0 ± 1.4 mV; L-AP5 = 7.7 ± 1.7 mV; control = 5.5 ± 1.2 mV; $P > 0.10$). Other strictly qualita-

Fig. 3 Group mean normalized slope-values for the 100 min experimental session during which low-frequency (0.05 Hz) test pulses were applied. The normalization was based on the initial 20-min baseline periods for each rat. High-frequency activation was applied 20 and 40 min after the start of each session. Insets show typical field potentials 5 and 55 min after the start of the session for each group. LTP was apparent in the control and L-AP5 animals only; half-decay times (calculated from log scores) were 162 and 45 min in these groups, respectively. All three groups show post-tetanic potentiation (p.t.p.) with half-decay times of 5.6, 1.9 and 4.6 mins for the control, D, L-AP5, and L-AP5 groups respectively, indicating that D, L-AP5 may have caused a diminution in the duration of p.t.p. However, our use of a 1 per 20 s test-pulse frequency was inadequate for analysing p.t.p. in detail. The first tetanus appears to have induced most (~85%) of the LTP, indicating that the stimulation parameters for inducing LTP were adequate. At the end of the experiment (60 min after the second tetanus), the slope values were 123.3%, 93.7% and 111.6% of the baseline mean for the control, D, L-AP5 and L-AP5 groups respectively.



tive observations indicated that paired-pulse facilitation and recurrent inhibition were also normal (not shown). However, D, L-AP5 caused a total blockade of the induction of LTP *in vivo* (Fig. 3). Infusion of L-AP5 brought about a faster decay of LTP than that in control rats but did not block its induction.

We have shown that AP5 impairs place but not visual discrimination learning—a profile similar, although not as severe, as that caused by hippocampal lesions (unpublished data and refs 7, 8). As place learning is relatively resistant to other selective disturbances of the brain including amygdala, thalamic and parietal lesions^{13–15}, to depletion of noradrenaline and striatal dopamine¹⁶, and as anticholinergic drugs impair both the spatial and visual tasks used here¹⁷, the profile of impairment with AP5 points to a mechanism involving interruption of hippocampal NMDA receptors. The mechanism remains a matter of speculation but could involve the voltage dependency of the ionic channels associated with NMDA receptors^{3,4}. Depolarization by one afferent input would allow a second input to cause an influx of Na⁺ and/or Ca²⁺ ions through the activated NMDA receptors, triggering long-term changes in synaptic efficacy, and giving rise to an 'association' between the two simultaneous inputs. Blockade of hippocampal NMDA receptors could, therefore, impair those types of learning mediated by hippocampal circuitry.

However, the demonstration that blockade of LTP by D, L-AP5 is associated with an AP5-induced learning impairment does not prove that these phenomena are causally related. Further experiments are required. However, we do now know that two very different antecedent treatments which block LTP—pharmacological receptor blockade and high-frequency stimula-

tion saturating synaptic enhancements¹⁸—are both sufficient to cause an anterograde spatial amnesia in the rat.

Finally, NMDA antagonists are known to be anticonvulsants¹⁹ and, subject to the development of similar compounds with adequate penetration of the brain²⁰, may soon be subject to clinical trials. It would be valuable to include thorough neuro-psychological testing of effects on memory within the design of these trials.

We thank Eric Harris, Frank Schottler and Mary English for advice and assistance. This work was supported by an MRC Fellowship (G8314974) to R.G.M.M.

Received 10 May; accepted 20 December 1985.

1. Watkins, J. C. & Evans, R. H. *A. Rev. Pharmac. Tox.* **21**, 165–204 (1981).
2. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462–465 (1984).
3. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 261–263 (1984).
4. Bliss, T. V. P. & Lomo, T. J. *J. Physiol., Lond.* **232**, 331–356 (1973).
5. Collingridge, G. L., Kehl, S. J. & McLennan, H. J. *J. Physiol., Lond.* **334**, 19–31 (1983).
6. Harris, E. W., Ganong, A. H. & Cotman, C. W. *Brain Res.* **323**, 132–137 (1984).
7. Morris, R. G. M., Garrud, P., Rawlins, J. N. P. & O'Keefe, J. *Nature* **297**, 681–683 (1982).
8. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Oxford University Press, 1978).
9. Andersen, P. *Prog. Brain Res.* **58**, 419–426 (1983).
10. McNaughton, B. L. in *Neurobiology of the Hippocampus* (ed. Siefert, W.) (Academic, London, 1983).
11. Lynch, G. & Baudry, M. *Science* **224**, 1057–1063 (1984).
12. Morris, R. G. M. *J. Neurosci. Meth.* **11**, 47–60 (1984).
13. Sutherland, R. J., Whishaw, I. Q. & Kolb, B. *Behavioural Brain Res.* **7**, 133–153 (1983).
14. Kolb, B., Pittman, K. & Sutherland, R. J. *Behav. Brain Res.* (in the press).
15. Kolb, B., Sutherland, R. J. & Whishaw, I. Q. *Behav. Neurosci.* **97**, 13–27 (1983).
16. Hagan, J. J., Alpert, J. E., Morris, R. G. M. & Iversen, S. D. *Behav. Brain Res.* **9**, 83–104 (1983).
17. Hagan, J. J., Tweedie, F. M. & Morris, R. G. M. *Behav. Neurosci.* (in the press).
18. Barnes, C. A. & McNaughton, B. L. *Soc. Neurosci. Abstr.* **9**, 191.16 (1983).
19. Croucher, M. J., Collins, J. F. & Meldrum, B. S. *Science* **216**, 899–901 (1982).
20. Meldrum, B. S. *Clin. Sci.* **68**, 113–122 (1985).

Elevation of intracellular calcium reduces voltage-dependent potassium conductance in human T cells

P. Bregestovski, A. Redkozubov* & A. Alexeev

USSR Cardiology Research Centre, Academy of Medical Sciences, Moscow, USSR

* Institute of Physiologically Active Substances, Academy of Sciences, Chernogolovka, Moscow Region, USSR

Both voltage-activated potassium channels^{1–4} and the concentration of free intracellular calcium^{5,6} have been implicated in the activation of T lymphocytes. Using the patch-clamp technique⁷, we now show an unexpected relationship between the level of intracellular calcium [Ca]_i in human lymphocytes and the amplitude of a voltage-dependent current: the elevation of [Ca]_i decreases the potassium conductance. This is in contrast to other systems where [Ca]_i activates K⁺ channels. Our results suggest that the level of intracellular calcium regulates the effective number of K⁺ channels capable of being activated.

We used the whole-cell configuration as the main mode of recording in our experiments and, on the basis of the observations of Marty and Neher⁸, we assume that, at least in steady-state conditions, the calcium concentration in the cell, [Ca]_i, corresponds to the concentration of free Ca²⁺ in the pipette. The voltage steps (given from a holding potential of –80 mV to a potential of more than –30 mV) produced outward currents that had kinetics and amplitudes similar to those described previously^{2,3,9} (Fig. 1a). These currents flow through K⁺-selective channels^{2,9}. They were blocked by 100 mM tetraethylammonium, 0.1 mM quinine and 5 mM cadmium. The threshold for the cells investigated was between –40 and –30 mV, and the conductance was maximal at around +40 mV. The recovery from inactivation was very slow, and so we had to apply the pulses at 1-min intervals.

The peak amplitude of the K⁺ current was strongly dependent on [Ca]_i (Fig. 1). For example, at a membrane potential of +20 mV, with [Ca]_i = 0.01 μM, the peak amplitude was 330 ± 224 pA and at [Ca]_i = 1 μM it was 72 ± 33 pA (mean ± s.d. for 10 or 11 cells). When the pipette was filled with solution containing 0.01 μM Ca²⁺, the amplitude of the outward currents usually increased during the first 5–10 min after the establishment of the whole-cell configuration, whereas when the pipette contained 1 μM Ca²⁺, the currents recorded decreased during the first 5–10 min after cell penetration.

The effect of [Ca]_i on the potassium conductance was confirmed in experiments with the Ca²⁺ ionophore A23187 (ref. 10), which is known to stimulate lymphocyte activation¹¹. As Fig. 2 shows, when 1 μM A23187 was added, the K⁺ inactivation was accelerated and the current amplitude decreased continuously. Ten minutes after A23187 application, the conductance was inhibited on average by 56% (4 cells). The presence of Ca-EGTA buffer inside the pipette might have been expected to prevent these [Ca]_i-induced changes. However, in similar conditions of low Ca buffering, the application of the ionophore leads to the activation of Ca²⁺-dependent K⁺ channels in the lacrimal gland¹². Thus, it is possible that in our experiments the reduced K⁺ conductance following the addition of A23187 is due to the elevation of [Ca]_i in the cell.

There are three main ways in which [Ca]_i could affect the potassium conductance: an alteration in single-channel conductance, an alteration in the kinetics of the channels or a change in the number of channels capable of being activated. To evaluate these possibilities, we examined the single-channel currents at different levels of [Ca]_i.

When the membrane potential is held at a depolarized level, the outward current is almost completely inactivated^{2,9} and the current can then be recorded through single channels in the whole-cell configuration (Fig. 3a). Figure 3b shows the voltage dependence of single-channel currents for several cells at two different [Ca]_i levels. The data have been averaged for four or five cells at each Ca²⁺ concentration. The slope conductance is ~14 pS when [Ca]_i is 0.01 μM and ~12 pS when [Ca]_i is 1 μM

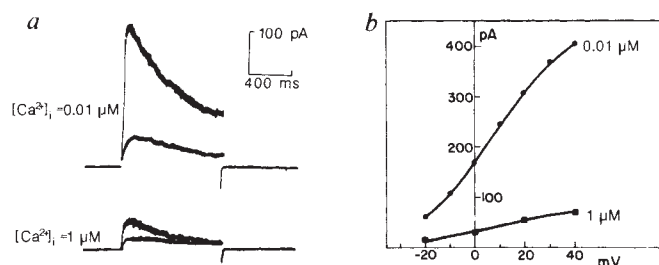


Fig. 1 Effect of $[Ca]_i$ on K^+ currents in human T lymphocytes. Whole-cell recording. *a*, Typical outward currents at $[Ca]_i = 0.01 \mu M$ (upper traces) and $[Ca]_i = 1 \mu M$ (lower traces). Depolarizing pulses of 50 and 100 mV were applied from a holding potential of -80 mV. *b*, Voltage dependence of outward currents at $0.01 \mu M$ and $1 \mu M$ $[Ca]_i$. Cells were of almost identical size (diameter $7 \pm 1 \mu m$). The linear leakage current was negligible compared with the voltage-dependent component and therefore was not subtracted. No analog compensation for series resistance was applied (the error was usually <5 mV). Peripheral blood mononuclear cells were separated from fresh heparinized blood from healthy donors by Ficoll-Hypaque density-gradient centrifugation and then resuspended in medium 199. The cells were initially depleted of macrophages by a single incubation ($37^\circ C$, 1 h) in plastic Petri dishes. T lymphocytes were purified as described elsewhere²¹ by rosetting macrophage-depleted peripheral blood mononuclear cells in medium 199 with unsensitized sheep erythrocytes and separating rosetted T lymphocytes from non-rosetted cells on Ficoll-Hypaque gradient. The pellet was then lysed with ammonium chloride to obtain T cells (93–97% reform E-rosettes). Before recording, the cells were kept in medium 199 at room temperature or $37^\circ C$. For patch-clamp experiments, the cells were placed in Petri dishes and maintained at room temperature. The bath solution contained (in mM): NaCl, 160; KCl, 4.5; $CaCl_2$, 2; $MgCl_2$, 1; glucose, 20; HEPES (pH 7.4), 10. The patch electrodes were filled with one of three solutions, containing (in mM): KCl, 140; $MgCl_2$, 2; HEPES pH 7.2, 10; EGTA, 1.1; and $CaCl_2$, 0.1 (pCa 8), or $CaCl_2$ 0.55 (pCa 7), or $CaCl_2$, 0.98 (pCa 6). In nine cells the Ca buffering was increased by using 11 mM EGTA and 1 mM $CaCl_2$ ($Ca^{2+} = 2$ nM, see ref. 2). No difference in G_{max} with data for the solution with pCa 8 was observed. The resistance of the pipettes ranged between 3 and 7 M Ω . The input resistance of the cells was 10–20 G Ω . Currents were recorded using an EPC-5 amplifier (List-Electronic) with built-in low-pass filter set to 1 kHz unless otherwise stated. Single-channel currents were digitized at a sampling interval of 0.4–1 ms and analysed with a microcomputer (Apple II).

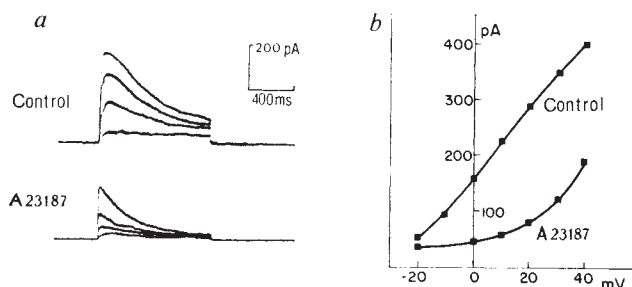


Fig. 2 Effect of the Ca^{2+} ionophore A23187 on K^+ currents. *a*, Responses to depolarizing voltage pulses of 60, 80, 100 and 120 mV before (upper traces) and 10 min after (lower traces) addition of $1 \mu M$ A23187. $0.01 \mu M$ Ca^{2+} in the pipette, holding potential -80 mV. Note the reduction in the current amplitudes and the acceleration of inactivation after treatment with A23187. Similar results were obtained with pipettes containing $0.1 \mu M$ Ca^{2+} . *b*, Dependence of K^+ currents on membrane potential before and after addition of $1 \mu M$ A23187. Same cell as in *a*. The unusual curvature of the lower curve is probably due to the continuous development of the ionophore effect. The depolarizing pulses were applied every minute, decreasing stepwise from $+40$ to -20 mV.

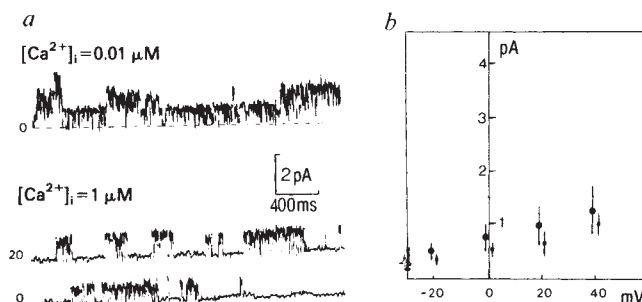


Fig. 3 Single-channel K^+ currents at different $[Ca]_i$ levels in human T-lymphocyte membrane. *a*, Currents recorded in the whole-cell configuration. The pipette contained $0.01 \mu M$ Ca^{2+} (upper traces) or $1 \mu M$ Ca^{2+} (lower traces). The membrane potential was held at the values indicated. 300-Hz filter. Note the higher frequency of channel activation at $0.1 \mu M$ $[Ca]_i$. *b*, Current-voltage relationships of K^+ channel at $0.1 \mu M$ (●) and $1 \mu M$ (■) $[Ca]_i$. The slope conductances are very similar (12–14 pS) for the two Ca concentrations. The values plotted are mean $\pm$ s.d. for four or five cells.

(Fig. 3b). Thus, the level of $[Ca]_i$ does not markedly affect the single-channel conductance. In some cells, channels with smaller conductance (6–8 pS) were also recorded (see also ref. 13). However, these appeared less frequently than the large-conductance channels and were observed at all levels of $[Ca]_i$. This type of channel was not analysed systematically.

The kinetic parameters of the 'large' K^+ channels were compared at two values of $[Ca]_i$, $0.01 \mu M$ and $1 \mu M$. The open time, which did not seem to depend on membrane potential, was not significantly affected by $[Ca]_i$, being 36 ms at low $[Ca]_i$ and 30 ms at high $[Ca]_i$. The fast component of the closed time (shut periods in the bursts) for both $[Ca]_i$ values was about 10 ms. The average 'burst time' was larger at low $[Ca]_i$ (250 ms) than at high $[Ca]_i$ (140 ms). However, this may have been due to the fact that at low $[Ca]_i$ the probability of opening was higher; the slow component of the closed time (periods between bursts) was 30–50 ms at low $[Ca]_i$ and 150–300 ms at high $[Ca]_i$. These results indicate that elevation of $[Ca]_i$ decreases either the probability of opening of the K^+ channels or the number of channels capable of being activated.

Experiments with inside-out patches provided direct evidence for the Ca-sensitivity of K^+ channels. Figure 4 shows that by increasing the bath Ca^{2+} concentration from $0.01 \mu M$ to $1 \mu M$, channel activity was almost completely blocked. In this patch

and in several others the blocking effect of Ca^{2+} was reversible. However, in many inside-out patches K^+ channels showed a tendency to degradation and, activity was not always recovered after washing out of high Ca^{2+} .

Our data differ from those of Cahalan *et al.*¹³, who failed to observe a Ca^{2+} -dependence of the K^+ conductance in human T cells¹³. We cannot explain this discrepancy. However, our results might explain an observation mentioned by these authors of an increase in peak current during the first few minutes after the transition to the whole-cell configuration. The intracellular Ca^{2+} concentration in lymphocytes is about $0.1 \mu M$ (ref. 5). Penetration of the cell with a pipette containing a low Ca^{2+} concentration (Cahalan *et al.*¹³ used pipettes containing 2 nM Ca^{2+} and fluoride as the main anion) could lead to an increase in the number of available K^+ channels and thus to an increase in the outward current. As mentioned above, with pipettes containing high $[Ca]_i$ we observed a decrease in the K^+ currents during the first few minutes after rupture of the membrane.

In conclusion, we suggest that the K^+ channels of the human T-lymphocytes membrane are Ca^{2+} -sensitive, but in an opposite way to that described for most other systems^{14–20}. It was shown recently that at relatively high $[Ca]_i$ ($\geq 0.1 \mu M$)²⁰ or at relatively

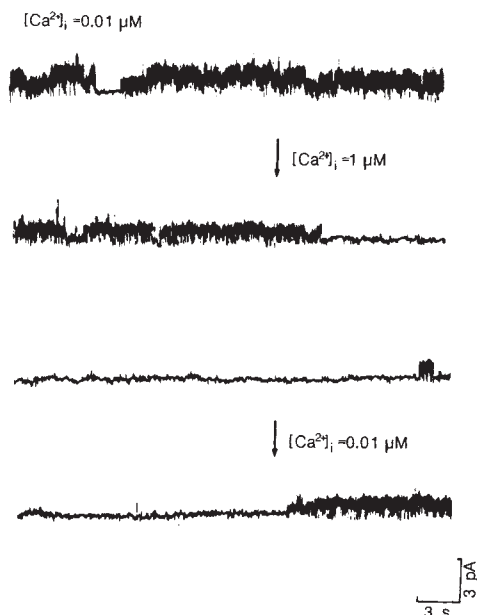


Fig. 4 Effect of Ca^{2+} on single-channel currents recorded from an inside-out membrane patch. The membrane potential was set at 40 mV. Symmetrical potassium solution (140 mM). The upper trace shows the channel activity with 0.01 μM Ca^{2+} in the bath. Note the presence of at least two channels in the patch. The next two traces show the inhibition of channel activity by 1 μM Ca^{2+} . The bottom trace shows the recovery of activity after washing with 0.01 μM Ca^{2+} . Sequential recording.

high membrane potentials (>40 mV)²⁰, the probability of opening of high-conductance Ca^{2+} -activated K^{+} channels decreases. We cannot exclude the possibility that this effect is present in our case. However, the blockade of the K^{+} conductance was observed at relatively low Ca^{2+} concentration ($[\text{Ca}]_i = 0.01 \mu\text{M}$) and at potentials of -30 to -20 mV. This process does not seem to be voltage-dependent: maximum conductance (G_{max}) was achieved at similar membrane potentials ($+20$ to $+50$ mV) for all cells investigated and usually did not decrease at higher potentials. For $[\text{Ca}]_i = 0.01 \mu\text{M}$, $0.1 \mu\text{M}$ and $1 \mu\text{M}$, G_{max} was, respectively, 3.7 ± 1.9 ($n = 19$), 1.8 ± 0.9 ($n = 5$) and 0.8 ± 0.5 ($n = 12$). Dividing G_{max} by the single-channel conductance, we estimate that at 0.01 μM $[\text{Ca}]_i$ the peak current is generated by a mean of about 280 channels; at 1 μM $[\text{Ca}]_i$, this number decreases to 60.

Kakei and Noma have observed a blocking effect of Ca^{2+} on ATP-sensitive K^{+} channels¹⁹. Our preliminary results show that K^{+} channels in T lymphocytes are not affected by intracellular ATP (up to 1 mM ATP in the pipette). Intracellular application of cyclic AMP (10 μM) also has no effect on voltage-dependent K^{+} conductance. The mechanism of channel regulation and the role of these channels in lymphocyte activation clearly require further investigation.

We thank Drs P. Ascher and A. Trautman for their comments on the manuscript.

Received 28 August; accepted 9 December 1985.

1. Fukushima, Y., Hagiwara, S. & Henkart, M. *J. Physiol., Lond.* **351**, 645–656 (1984).
2. DeCoursey, T. E., Chandy, K. G., Gupta, S. & Cahalan, M. D. *Nature* **307**, 465–468 (1984).
3. Matteson, D. R. & Deutsch, C. *Nature* **307**, 468–471 (1984).
4. Chandy, K. G., DeCoursey, T. E., Cahalan, M. D., McLaughlin, C. & Gupta, S. *J. exp. Med.* **160**, 369–385 (1984).
5. Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68–70 (1982).
6. Ashman, R. F. in *Fundamental Immunology* 267–300 (Plenum, New York, 1984).
7. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
8. Marty, A. & Neher, E. in *Single Channel Recording* (eds Sakmann, B. & Neher, E.) Ch. 7 (Plenum, New York, 1983).
9. Bregestovski, P., Redkozubov, A., Krasnikova, T. & Radyukhin, V. *Biol. Membr.* **2**, 319–322 (1985) (in Russian).

10. Selinger, Z., Eimeri, S. & Schramm, M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 128–131 (1974).
11. Maino, V. G., Green, N. M. & Creighton, M. J. *Nature* **251**, 324–326 (1974).
12. Marty, A., Tan, Y. & Trautmann, A. *J. Physiol., Lond.* **357**, 293–325 (1984).
13. Cahalan, M. D., Chandy, K. G., DeCoursey, T. E. & Gupta, S. *J. Physiol., Lond.* **358**, 197–237 (1985).
14. Marty, A. *Nature* **291**, 497–500 (1981).
15. Lux, H. D., Neher, E. & Marty, A. *Pflügers Arch. ges. Physiol.* **389**, 293–295 (1981).
16. Siegelbaum, S. A., Camardo, J. S. & Kandel, E. R. *Nature* **299**, 413–417 (1982).
17. Barrett, J., Magleby, K. & Palotta, B. *J. Physiol., Lond.* **331**, 211–230 (1982).
18. Vergara, C. & Latorre, R. *J. gen. Physiol.* **82**, 543–568 (1983).
19. Kakei, M. & Noma, A. *J. Physiol., Lond.* **352**, 265–284 (1984).
20. Findly, I., Dunne, M. J. L. & Petersen, O. H. *J. Membrane Biol.* **83**, 169–175 (1985).
21. Gupta, S. *J. Immun.* **123**, 2664–2668 (1979).

Is the T-cell receptor involved in T-cell killing?

Antonio Lanzavecchia

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4005 Basel, Switzerland

It is known that when two populations of cytotoxic T lymphocytes (CTL) are mixed in conditions where antigen recognition can occur in only one direction, killing also proceeds only in the same direction^{1,2}. These data suggest that occupancy of the T-cell receptor (TCR) is required for the expression of the lytic function by effector CTL, but do not establish whether the TCR itself has a role in the killing process. In particular, it is not clear whether the TCR is involved in the actual delivery of the lethal hit to the target cell (either being itself part of the lytic machinery or directing it), or whether TCR occupancy only serves the function of triggering a set of lytic reactions which are themselves non-specific and not directed by the TCR. The use of mitogenic lectins or mitogenic antibodies, which bypass specific recognition and induce nonspecific killing^{3–5}, also does not help to clarify this issue, since a necessary characteristic of these ligands is that they bind to the TCR complex or to other 'triggering' molecules and probably bridge these structures to the target cell^{6,7}. The present study describes an *in vitro* system using human T-cell clones which allows us to dissociate the triggering of a CTL from the delivery of the lethal hit, using no externally added ligands. We report that, once triggered by recognition of the specific target, a CTL can kill any other cell that binds to it, indicating that TCR occupancy is required for triggering, but not for the delivery of the lethal hit.

The experimental design and the cell clones used are shown in Fig. 1. The question is: does a CTL clone (cell 1), once triggered by its specific target (cell 2), become able to kill a T cell that is specific for the alloantigens of the CTL itself (cell 3), and is this killing limited to cells that specifically bind to the CTL or is it also effective on bystander cells, or cells that are in close proximity to the activated CTL because they bind to the same target (cells 4, 5 and 6)?

Figure 2 shows a representative experiment and analyses the lytic activity of one CTL clone bearing the T8 antigen. Figure 2A shows the killing of the appropriate 'triggering' target (cell 2), while Fig. 2B, C show that two distinct T-helper clones specific for alloantigens of CTL 1 (cell 3) are not killed when incubated with the CTL alone or with the CTL in the presence of an inappropriate triggering target cell. However, if the appropriate triggering target for the CTL (cell 2) is added, cell 3 is also killed. We refer to this killing as backwards killing, since in this case the CTL cannot recognize cell 3 through the TCR, whereas cell 3 does recognize and bind to the CTL. Backwards killing, although always less efficient than direct killing, was rather effective, ranging from 35 to 80% specific release in 20 different tests using 5 different CTL clones and 4 different alloreactive helper clones. Furthermore, the efficiency of backwards killing was always strictly proportional to the efficiency of direct killing and, in addition, no backwards killing

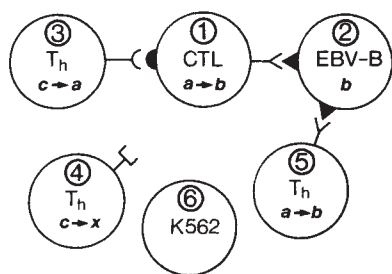


Fig. 1 Experimental rationale. The addition of the 'triggering' target (cell 2) leads to activation of a CTL (cell 1) and killing of cell 2. In these conditions, when the TCR is occupied by cell 2, is killing directed only against cell 2 or does CTL 1 also become capable of lysing other cells which it does not specifically recognize? For example cell 3, alloreactive T-helper clones recognizing CTL 1; cell 4: T-helper clones specific for neither cell 1 nor cell 2; cell 5, alloreactive T-helper clones recognizing cell 2 and therefore in close proximity to cell 1; cell 6, NK-sensitive target K562. **Methods.** Alloreactive T-cell clones and EBV-transformed B cells were obtained as described previously¹². The specificity and function of helper and CTL clones were defined on the basis of: (1) proliferative response to irradiated peripheral blood mononuclear cells and EBV-B cells, (2) killing of phytohaemagglutinin (PHA)-activated blasts and EBV-B cells and (3) help for polyclonal activation of peripheral blood B cells¹³. The T cells used in these experiments were obtained from interleukin-2-dependent cultures 3–6 weeks after the last antigenic stimulation. The following clones were used: Cell 1, five alloreactive CTL clones obtained from a mixed lymphocyte reaction (MLR) from donor *a* (HLA-A2, Aw26, Bw58, Bw49, DC1, DC6) against *b* (A1, A3, B8, Bw60, Cw3, Cw7, DR3, DR5); they kill cells of HLA type *b* but not *c*-type (Aw30, A32, Bw38, B39, DC1, DR6, DR8) cells. Four are T8⁺ and one is T4⁺. Cell 2, EBV-transformed B cells from donor *b*. Cell 3, four alloreactive T4⁺ T-helper clones from donor *c*, anti-*a*. They proliferate in response to *a*-type but not *b*-type cells. Cell 4, alloreactive T-helper clones or PHA blasts from donor *c* not reactive to either *a* or *b*. Cell 5, alloreactive T-helper clones from an MLR *a* anti-*b* but not reactive against *c*.

was observed when two alloreactive helper clones were used as cell 1 (data not shown). This suggests that backwards killing is a property of antigen-triggered CTL.

A slightly different question is whether cells that do not bind specifically to CTL 1 can also be killed. Figure 2D–F shows that bystander T cells (cell 4) or a T-cell clone reactive with cell 2 and therefore in close proximity to CTL 1 (cell 5) are lysed only slightly and here again only in conditions of optimal triggering of the CTL. This can be defined as bystander killing, since in this case there is no specific interaction between the CTL and the cell that is killed. The bystander killing was rather variable and always less efficient than backwards killing, ranging between 0 and 20% of the backwards killing in 35 different tests. It is possible that bystander killing is mediated by the 'leakage' of soluble lytic factors^{8,9} which would also be active (although with much lower efficiency) on cells that are not in direct contact with a triggered CTL. However, since it is known that a low degree of conjugate formation occurs in the absence of specific recognition², we favour the hypothesis that cell contact is also required and responsible for the bystander killing.

An old question is whether the CTL kills itself in the process of recognizing and killing other cells. Figure 2G shows that this is not the case, because a very low degree of auto-killing is observed only in the presence of an excess of activated unlabelled CTL and is not higher than that observed for bystander T cells. Therefore, although the backwards killing is not mediated through the TCR, it is still directional in a sense, since the effector CTL does not kill itself.

In this clonal system, none of the five CTL clones tested ever acquired lytic activity against the natural killer (NK)-sensitive target K562, even when simultaneously triggered by the

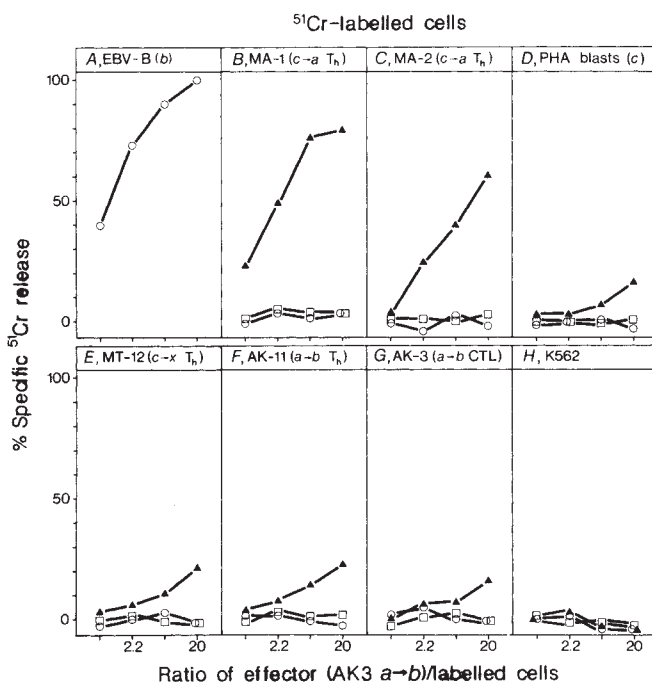


Fig. 2 Antigen-triggered CTL can kill backwards. Cells from a T8⁺ CTL clone (AK-3, specificity *a* anti-*b*) were incubated in 200 μ l RPMI-10% fetal calf serum in V-bottom plates with 5×10^3 ⁵¹Cr-labelled target cells at different effector:target ratios. Each panel represents the specific lysis of the ⁵¹Cr-labelled cells indicated. These labelled cells were incubated: with CTL alone (○), with the CTL plus 2×10^4 EBV-B cells type *b*, that is, the triggering target (▲), or with 2×10^4 EBV-B cells type *c*, that is, an inappropriate target (□). ⁵¹Cr-labelled cells were: A, EBV-B cells type *b* (cell 2); B, C, alloreactive helper clones *c* anti-*a* (cell 3); D, E, PHA blasts or helper clone from donor *c* unreactive to both *a* and *b* (cell 4); F, alloreactive helper clone *a* anti-*b* (cell 5); G, CTL clone AK-3 (cell 1); H, K562 (cell 6). After incubation at 37 °C, for 8h, 100 μ l of supernatant was collected and the specific release was determined as described elsewhere¹⁴. Spontaneous release never exceeded 30%.

appropriate target (Fig. 2H). This finding indicates that NK activity cannot simply be related to the 'activated' state of a CTL, but instead that some specific interaction may take place between NK cells and their targets¹⁰.

The above experiments show that once a CTL is triggered by recognizing its specific target, the lytic activity can be directed against cells that are not recognized by the TCR but which are themselves binding to, and therefore in close contact with, the triggered CTL. This demonstrates that the TCR is required for the triggering of the CTL effector function but not for the delivery of the lethal hit.

If cell 3 is actually committing suicide by binding to a triggered CTL, it ought to be possible to prevent its death by diverting it away from this lethal interaction. One way to create such a diversion is to offer cell 3 a non-cytolytic cell carrying the same alloantigens as CTL 1. Figure 3 shows that the addition of Epstein-Barr virus (EBV)-transformed B cells autologous to the CTL is an extremely effective way of averting the death of cell 3, while third-party B cells provide no significant protection. Furthermore, it is possible to inhibit allorecognition by cell 3 (T4⁺) using an anti-T4A monoclonal antibody, which itself is not affecting antigen recognition by CTL 1 (T8⁺). As shown in Fig. 3, addition of anti-T4A antibody also effectively inhibits the killing of cell 3, while the killing of the appropriate target is not affected (data not shown). From these experiments we conclude that cell contact is required for backwards killing and that the 'specificity' of backwards killing is identical to the

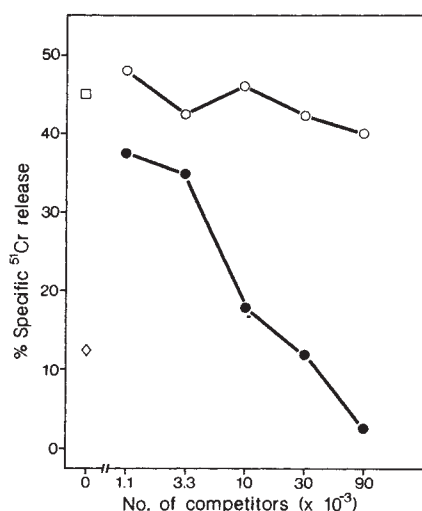


Fig. 3 Specificity of backwards killing is that of the cell which is killed. 5×10^3 ^{51}Cr -labelled cells from a T4⁺ T-helper clone (MA-2, c anti-a) were incubated with 5×10^4 cells from an a anti-b CTL clone (AK-5) and 2×10^4 b-type EBV-B cells. Increasing numbers of EBV-B cells were added to inhibit backwards killing: ●, EBV-B cells type a; ○, type c. ◇ Indicates the level of killing obtained in the presence of $10 \mu\text{g ml}^{-1}$ anti-T4A monoclonal antibody (provided by E. Reinherz). □ Indicates killing in the absence of any competitors.

specificity of the cell that is killed rather than the specificity of the killer. Similar properties are described by Rammensee *et al.*¹¹ for mature CTL, which can prevent the activation of T-killer precursors directed against them. We predict that if backwards killing is the mechanism of such 'veto' activity, it will correlate directly with the degree to which the CTL are triggered by the correct target.

Two particularly critical technical aspects to the demonstration of backwards killing require some comment. First, backwards killing is better shown at 8 than at 4 hours. This is not unexpected, since this type of killing is a three-cell interaction and would therefore occur less frequently than the two-cell interaction of the CTL and its target. Second, since backwards killing is dependent on CTL triggering, the specific 'triggering' targets should be in excess. In fact, in initial experiments, when T-cell clones were used as a source of 'triggering' targets (cell 2), the level of backwards killing observed was low and variable. Higher and reproducible levels of backwards killing were obtained when EBV-B cells were used as 'triggering' targets. These cells express high levels of major histocompatibility complex class I and class II antigens and are, on a per cell basis, 10–100 times more efficient than T-cell blasts in triggering proliferation of CTL clones (data not shown).

The above experiments allow the dissociation of the triggering step from the delivery of the lethal hit and point to the following conclusions: (1) mature CTL are not constitutively cytotoxic but become cytotoxic only when the TCR is occupied by the 'triggering' antigen, (2) the cytotoxic activity is not mediated through the TCR nor directed by it, since a CTL, once triggered, can also kill cells that are not recognized by the TCR itself.

I thank Doris Scheidegger for technical assistance, Judie Hossmann for preparation of the manuscript, Hans-Georg Rammensee and Franco Celada for reading the manuscript and Polly Matzinger for critical comments throughout this work. The Basel Institute for Immunology was founded and is supported by F. Hoffman-La Roche & Co., Basel, Switzerland.

Received 15 November 1985; accepted 6 January 1986.

1. Kuppers, R. C. & Henney, C. S. *J. exp. Med.* **143**, 684–689 (1976).
2. Fishelson, Z. & Berke, G. *J. Immun.* **120**, 1121–1126 (1978).
3. Bevan, M. J. & Cohn, M. *J. Immun.* **114**, 559–565 (1975).

4. Leeuwenberg, J., Spits, M., Tax, W. & Capel, P. *J. Immun.* **134**, 3770–3775 (1985).
5. Siliciano, R. F., Pratt, J. C., Schmidt, R. E., Ritz, J. & Reinherz, E. L. *Nature* **317**, 429–430 (1985).
6. Chilson, O. P., Boylston, A. W. & Crumpton, M. J. *EMBO J.* **3**, 3239–3245 (1984).
7. O'Flynn, K., Krensky, A. M., Beverly, P. C. L., Burakoff, J. T. & Linch, D. C. *Nature* **313**, 686–687 (1985).
8. Dennert, G. & Podack, E. R. *J. exp. Med.* **157**, 1483–1495 (1983).
9. Henkart, P. A., Millard, P. J., Reynolds, C. W. & Henkart, M. P. *J. exp. Med.* **160**, 75–93 (1984).
10. Hercend, T. *et al. Eur. J. Immun.* **14**, 844–852 (1984).
11. Rammensee, H. G., Bevan, M. J. & Fink, P. J. *Immun. Today* **6**, 41–43 (1985).
12. Lanzavecchia, A. *Nature* **314**, 537–539 (1985).
13. Lanzavecchia, A., Parodi, B. & Celada, F. *Eur. J. Immun.* **13**, 733–738 (1983).
14. Brunner, K. T., Muel, J., Cerottini, J. C. & Chapuis, B. *Immunology* **14**, 181–196 (1968).

Differential expression of *myc* family genes during murine development

Kathryn A. Zimmerman, George D. Yancopoulos, Robert G. Collum, Russell K. Smith, Nancy E. Kohl, Kathleen A. Denis*, Marion M. Nau†, Owen N. Witte*, Dominique Toran-Allerand‡, Connie E. Gee§, John D. Minna† & Frederick W. Alt

Departments of Biochemistry and ‡Center for Reproductive Sciences, College of Physicians and Surgeons of Columbia University, New York, New York 10032, USA

* Department of Microbiology and Molecular Biology Institute, University of California, Los Angeles, California 90024, USA

† NCI-Navy Medical Oncology Branch, National Cancer Institute, National Institutes of Health & Naval Hospital, Bethesda, Maryland 20814, USA

§ Dana Farber Cancer Institute, Harvard University Medical Center, Boston, Massachusetts 02115, USA

The *myc* family of cellular oncogenes contains three known members. The *N-myc* and *c-myc* genes have 5'-noncoding exons, strikingly homologous coding regions^{1–3}, and display similar oncogenic potential in an *in vitro* transformation assay^{4,5}. The *L-myc* gene is less well characterized, but shows homology to *N-myc* and *c-myc* (ref. 6; also see below). *c-myc* is expressed in most dividing cells, and deregulated expression of this gene has been implicated in the development of many classes of tumours⁷. In contrast, expression of *N-myc* has been found only in a restricted set of tumours, most of which show neural characteristics; these include human neuroblastoma^{8–10}, retinoblastoma^{10,11} and small cell lung carcinoma (SCLC)¹². *L-myc* expression has so far been found only in SCLC⁶. Activated *N-myc* and *L-myc* expression has been implicated in oncogenesis^{4–6,11,12}; for example, although *N-myc* expression has been found in all neuroblastomas tested^{9,10}, activated (greatly increased) *N-myc* expression, resulting from gene amplification^{8,13,14}, is correlated with progression of the tumour¹⁵. We now report that high-level expression of *N-* and *L-myc* is very restricted with respect to tissue and stage in the developing mouse, while that of *c-myc* is more generalized. Furthermore, we demonstrate that *N-myc* is not simply a neuroectoderm-specific gene; both *N-* and *L-myc* seem to be involved in the early stages of multiple differentiation pathways. Our findings suggest that differential *myc* gene expression has a role in mammalian development and that the normal expression patterns of these genes generally predict the types of tumours in which they are expressed or activated.

To assay for tissue- or stage-specific *myc* expression, we first isolated specific probes for the murine *N-* and *L-myc* genes (Fig. 1 legend). Steady-state expression levels were determined by assaying total or poly(A)⁺ RNA using electrophoresis/blotting procedures for hybridization to *c-*, *N-* or *L-myc* probes; typical results are shown in Fig. 1. The *N-myc* probe identified the murine *N-myc* messenger RNA as a 2.9-kilobase (kb) transcript in a mouse neuroblastoma cell line (NBA2) and in certain other tissues (Fig. 1), but not in many other types of murine cell lines

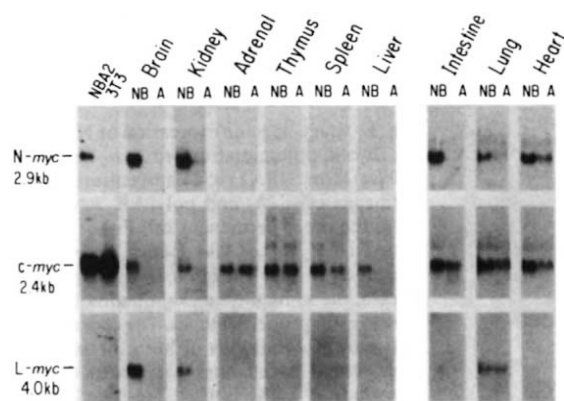


Fig. 1 Expression of *N-myc*, *L-myc* and *c-myc* in murine newborn and adult tissues; representative Northern blotting analysis. Approximately 5 μ g of poly(A)⁺ RNA from the NBA2 and 3T3 cell lines and 20 μ g of total RNA from various tissue samples isolated from newborn (NB) or adult (A) mice was fractionated by formaldehyde/agarose gel electrophoresis, transferred to nitrocellulose and assayed for hybridization to a ³²P-labelled murine *N-myc* (upper), *c-myc* (middle), or *L-myc* (lower) probes as described previously¹⁰. The filter was sequentially hybridized to the three different probes, the signal from the previous probe was removed before the next probe was applied¹⁰. Sources of the RNA are indicated above the corresponding panels; the brain panel indicates the results obtained with RNA from the forebrain (olfactory bulb, cerebral cortex, hippocampus, neostriatum, septum, hypothalamus, thalamus and amygdala). The three panels on the right were from an independent filter; exposure relative to the six panels on the left (standardized to the brain signal) was approximately three times longer for *N-* and *L-myc* and five times longer for *c-myc* in order to visualize the signals more easily.

Methods. The *N-myc*-specific cDNA clone was isolated by screening a cDNA library prepared from a murine pre-B-cell line²³ with a human *N-myc* cDNA clone (pI1-1)¹. The isolated murine cDNA clone (pM.2) represents 550 base pairs (bp) of the 3'-untranslated region of the murine *N-myc* gene². A panel of murine *myc* genes was isolated from a genomic *EcoRI* library (in Charon 30) on the basis of their hybridization to both second and third exon probes isolated from the human *N-myc* gene (K.A.Z. *et al.*, in preparation). One of these clones, a 12-kb murine genomic *EcoRI* fragment, was demonstrated to have strong homology to the *SmaI/EcoRI* probe for the human *L-myc* gene⁶ and weak homology to *N-myc* or *c-myc* probes. From this murine genomic clone, an 800-bp *SacI* fragment (murine *L-myc* probe) with strong homology to the human *L-myc* probe was isolated. Various criteria indicated that the murine *L-myc* probe was derived from the murine homologue of the human *L-myc* gene (K.A.Z. *et al.*, in preparation). The *c-myc* probe was the internal 1-kb *XhoI* fragment of the *c-myc* cDNA clone, pMcm54 (ref. 24). The *myc* probes did not cross-hybridize with other *myc* gene sequences in the hybridization conditions used here.

(for example, 3T3 cells (Fig. 1), but see below). The NBA2 line has not amplified the *N-myc* gene and, correspondingly, appeared to express *N-myc* mRNA at levels similar to those found in human neuroblastomas, in which the gene is not amplified (ref. 10 and K.A.Z. *et al.* data not shown). The *L-myc* probe identified a ~4.0-kb RNA sequence in certain murine tissues (Fig. 1, Table 1), but did not hybridize to RNA from any murine cell line tested (Fig. 1 and see below).

For quantitative studies, multiple blotting experiments with independent RNA preparations from tissues of the same and different mouse strains were performed. The relative steady-state level of a particular *myc* mRNA was obtained by averaging values obtained from densitometric analyses of independent experiments and expressed as a percentage of its level in the tissue in which expression was highest (Table 1; see legend for details). Solution hybridization analyses confirmed the blotting results and demonstrated that the absolute levels of *N-myc* RNA

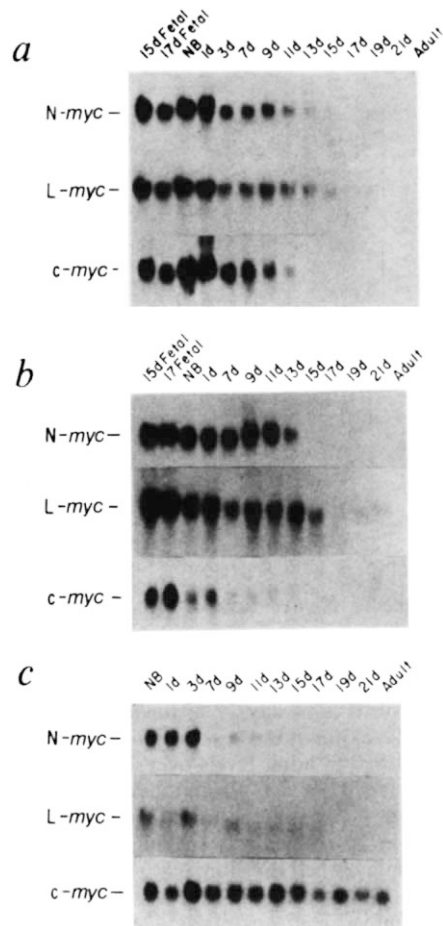


Fig. 2 Stage-specific expression of *myc* family genes. Total RNA was prepared from forebrain (a), hindbrain (midbrain, pons, cerebellum and medulla) (b) and kidney (c) of Columbia RIII mice at the indicated pre- or postnatal days and analysed for *N-myc*, *L-myc* and *c-myc* expression as described in Fig. 1 legend. Each lane contains 20 μ g of 'total RNA'.

in tissues and unamplified cell lines are very low ($\leq 0.005\%$ of poly(A)⁺ mRNA in the most abundant murine sources; data not shown). In normal tissues, expression of *N-myc* mRNA was greatest in newborn forebrain, hindbrain, kidney and intestine; strikingly, expression in these tissues was at least as great as that observed in the NBA2 neuroblastoma line (Fig. 1, Table 1). *N-myc* expression was detected in all other newborn tissues, but levels were, in most cases, more than 20-fold less than those of newborn brain (Fig. 1, Table 1). In the adult forebrain, hindbrain, kidney and intestine, *N-myc* expression was at least 20-fold less than in the corresponding newborn tissues, and in tissues where *N-myc* expression was lower at birth, a substantial decrease in adult expression was observed only in liver (Fig. 1, Table 1). Expression patterns of *L-myc* in tissues were similar to those of *N-myc*. Its expression was highest in newborn forebrain, hindbrain and kidney, although the level in kidney relative to that in brain was much lower than observed for *N-myc* (Fig. 1, Table 1). *L-myc* was detected at lower levels in lung and intestine and was not detectable in other tissues examined (Fig. 1, Table 1). As with *N-myc*, there was a dramatic decrease in *L-myc* expression between the newborn and adult forebrain, hindbrain and kidney; *L-myc* expression did not decrease in adult lung (Fig. 1, Table 1). In contrast to *N-myc* and *L-myc*, the relative level of the *c-myc* transcript varied no more than threefold among most newborn tissues; a decrease in *c-myc* expression between the newborn and adult was noted in tissues

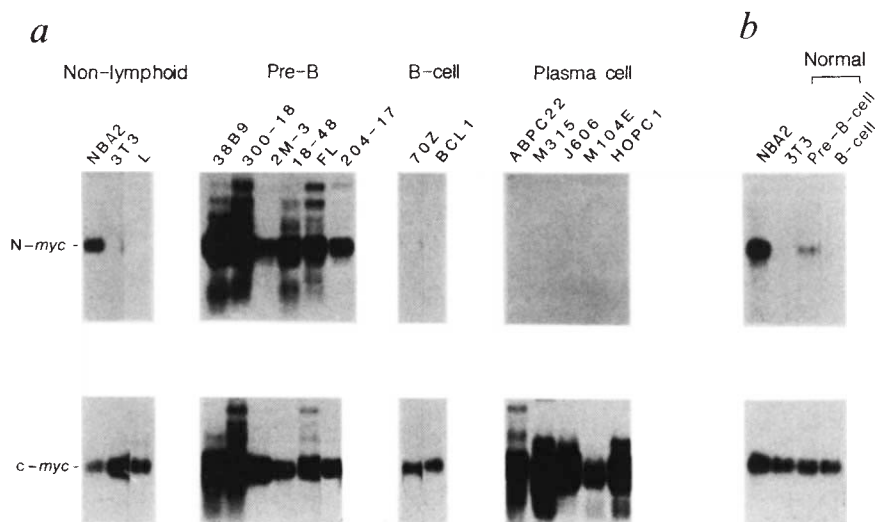


Fig. 3 Stage-specific expression of *N-myc* during B-cell differentiation *in vitro*. Total RNA (10 µg) from NBA2 (*a, b*) and cultured normal B-lineage cells (*b*) and poly(A)⁺ RNA (5 µg) from the other permanent cell lines (*a, b*) was assayed for *N*-, *c*- and *L-myc* expression as described in Fig. 1 legend. In the experiments shown, the same filter was sequentially hybridized to each probe. **Methods.** The characteristics and growth conditions of both the permanent and normal B-cell lines have been described elsewhere^{17,25}. The normal pre-B-cell line was clone 1 and the B-cell line was clone 2¹⁷.

where a decline in *N-myc* (and *L-myc*) expression occurred (Fig. 1, Table 1).

To elucidate the developmental expression of *myc* family genes further, we analysed forebrain, hindbrain and kidney from day 15 before birth until day 21 of postnatal development. In each tissue, *myc* expression was constant on all prenatal days tested and declined to adult levels over the first weeks of postnatal development (Fig. 2*a-c*). Significantly, the time course for the decline of *c-myc* differed from that of *N-myc* and *L-myc*, the difference being particularly striking for kidney. Notably, hypertrophy replaces cell division as a major means of growth prior to or coincident with the decrease in *myc* expression in the organs examined, and organ size does not increase enough to account for the observed decreases. Therefore, it seems likely that *myc* expression decreases within differentiating cells of a given lineage in brain or kidney rather than by mechanisms such as overgrowth of *myc*-producing cells by non-*myc*-producing cells.

The finding of at least low-level *N-myc* expression in all newborn tissues is surprising because this gene is not expressed in most tumours and cell lines. Low-level *N-myc* expression in tissues could result from high-level expression in specific subpopulations of cells within the tissue. For example, the difference in *N-myc* expression between the newborn and adult liver could be related to the presence of a differentiating haematopoietic cell population in this organ in the fetus (neonate) but not the

adult. To test this possibility, we assayed *N-myc* mRNA levels in permanent and normal cell lines representing multiple stages of the B-cell differentiation pathway. Significantly, *N-myc* was expressed in all of 17 Abelson murine leukaemia virus (A-MuLV) transformed pre-B-cell lines and not in any lines representing more mature B-cell stages (Fig. 3; representative data are shown). Quantitation experiments indicated that *N-myc* expression in pre-B-cell lines was similar to or in some cases significantly lower than those of neuroblastoma lines in which the *N-myc* gene was not amplified (data not shown). *N-myc* expression was not detectable in A-MuLV-transformed 3T3 cells (data not shown) or in an A-MuLV-transformed plasmacytoma (ABPC22; Fig. 3, lane 12), suggesting that A-MuLV transformation itself is not responsible for *N-myc* expression. In contrast to *N-myc*, *c-myc* expression was detected in all pre-B- and B-cell lines tested (Fig. 3*a, b*), while *L-myc* expression was never detected (data not shown). Although it has recently been reported that A-MuLV transformation of fibroblast cell lines is associated with *c-myc* gene amplification¹⁶, neither *c*- nor *N-myc* was amplified in the A-MuLV-transformed pre-B-cell lines (data not shown).

Additional evidence for the stage-specific expression of *N-myc* in pre-B cells was provided by examination of two cell lines clonally isolated from long-term cultured murine bone marrow¹⁷; the immunoglobulin gene rearrangement and expression patterns of these lines indicated that one represented the pre-B

Table 1 Relative *myc* RNA expression as a percentage of highest expressing tissue

	Cell lines		Forebrain	Hindbrain	Kidney	Adrenal	Thymus	Spleen	Liver	Intestine	Lung	Heart
	NBA2	3T3	NB/A	NB/A	NB/A	NB/A	NB/A	NB/A	NB/A	NB/A	NB/A	NB/A
<i>N-myc</i>	45	*	100/<5	80/<5	95/<5	5/<5	<5/<5	<5/<5	<5/*	30/<5	<5/<5	10/<5
	(2)	(2)	(3) (3)	(3) (3)	(3) (3)	(2) (2)	(2) (2)	(2) (2)	(2) (2)	(1) (1)	(1) (1)	(2) (2)
<i>c-myc</i>	300	300	60/<5	60/<5	35/10	45/70	100/70	100/35	90/<5	40/20	25/<5	15/5
	(2)	(2)	(3) (3)	(3) (3)	(3) (3)	(2) (2)	(2) (2)	(2) (2)	(2) (2)	(1) (1)	(1) (1)	(2) (2)
<i>L-myc</i>	*	*	100/*	100/*	10/*	*/	*/	*/	*/	<5/<5	5/5	*/
	(2)	(2)	(1) (1)	(1) (1)	(1) (1)	(1) (1)	(1) (1)	(1) (1)	(1) (1)	(1) (1)	(1) (1)	(1) (1)

The relative expression level of *myc* genes in tissues of Columbia RIII mice was determined as follows. Aliquots (20 µg) of total RNA from the various sources were assayed for hybridization to a given *myc* gene probe by the electrophoresis/blotting assay described in Fig. 1 legend, and the relative intensity of the *myc*-specific mRNA bands on the resulting autoradiograms was determined by densitometric scanning (using the Joyce Loeb Chromoscan 3). For each independent experiment, the intensity of a particular *myc* mRNA band in a given tissue was calculated as a percentage of the signal intensity generated with RNA from the most strongly expressing tissue (newborn forebrain for *N*- and *L-myc* and newborn thymus for *c-myc*), which was arbitrarily set at 100%. Each value given is an average (rounded off to the nearest multiple of five) of from one to three independent experiments (number in parentheses), each of which used an independent preparation of RNA. Similar results were obtained with total and poly(A)⁺ RNA from tissue samples of BALB/c mice (not included in the calculations). A value of <5 indicates that expression of the sequence in that tissue was detectable, although clearly below the 5% level and too low to obtain an accurate reading. An asterisk indicates that expression of the sequence was not detectable in that tissue and thus was far below the 5% level. NB, newborn; A, adult.

stage while the other represented the B-cell stage. Expression of N-myc mRNA was readily detectable in the pre-B but not in the B-cell cultures (Fig. 3b), while expression of c-myc occurred at similar levels in both cultures (Fig. 3b). Again L-myc expression was not detected in either culture (data not shown).

Our findings demonstrate that myc family genes display distinct tissue- and stage-specific expression patterns. The N-myc gene is clearly expressed in a stage-specific manner in differentiating B-lymphoid cells and expression of the N- and c-myc genes is differentially regulated during progression of this lineage (Fig. 3). In addition, the myc genes are differentially expressed during postnatal development of the kidney and probably during development of the brain (Fig. 2). Our findings demonstrate that N-myc and L-myc may have early stage-specific roles in multiple differentiation pathways. Because L-myc was identified from a large pool of unique murine myc genes originally isolated on the basis of their homology to both the second and third exons of the human N-myc gene (see Fig. 1 legend), it seems likely that the myc gene family contains additional members. Our findings raise the possibility that differential, or perhaps combinatorial, expression of these genes may be important in normal development.

We have suggested that activated myc gene expression in certain tumours may be mediated by expression of these genes in normal progenitor cells⁴. Our current findings support this notion. The finding of similar levels of c-myc expression in most embryonic and neonatal organs is consistent with the general expression and frequent activation (by a variety of mechanisms) of this gene in a wide spectrum of tumours. Activated expression of N-myc in embryonic tumours of neural origin, as well as activation of N- or L-myc in many SCLC (which have 'neural' characteristics^{18,19}), correlates with higher expression of these genes in developing brain and newborn intestine (which has a large population of developing neurones). The expression of L-myc in both the neonatal and adult lung is intriguing with respect to its potential relationship to SCLC; the pulmonary endocrine cells that give rise to this cancer are abundant in the fetal lung and persist in reduced numbers in the adult lung²⁰. The fetal and newborn kidney was the only non-neural organ in which relatively high-level expression of N-myc was detected. Guided by this result, we have recently found N-myc expression in Wilms' tumour²¹, a renal neoplasm which, again, arises from embryonic cells²².

Various criteria indicate that myc genes are involved in the genesis or progression of tumours in which their expression is activated^{4,5,11,15}. Whether lower-level expression of a particular myc gene, such as in pre-B-cell tumours or some unamplified neuroblastomas, contributes directly to tumorigenesis or simply reflects the differentiation state of the cell remains unknown. However, our findings suggest that activated N-myc and possibly L-myc expression is involved in other cancers that arise from primitive normal cell precursors.

We thank K. B. Marcu for the pMcm54 clone. This work was supported by NIH grants 2-PO1 CA 23767 and AI20047 (F.W.A.) and HDO834 (D.T.-A.), ACS grant CD-269 (F.W.A.), a Whitehall Foundation grant (D.T.-A.) and a Searle Scholars Award (F.W.A.). K.A.Z. was supported by NIH training grant CA 09503 and C.E.G. by an NIH NRSA fellowship. D.T.-A. is recipient of an NIMH research scientist development award (MH-00192) and F.W.A. is an Irma T. Hirsch career scientist and a Mallinckrodt Scholar.

10. Kohl, N. E., Gee, C. E. & Alt, F. W. *Science* **226**, 1335-1336 (1984).
11. Lee, W. H., Murphee, A. L. & Benedict, W. F. *Nature* **309**, 458-460 (1984).
12. Nau, M. M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
13. Schwab, M. *et al. Nature* **305**, 245-248 (1983).
14. Michitsch, R. W., Montgomery, K. T. & Melera, P. W. *Molec. cell. Biol.* **4**, 2370-2380 (1984).
15. Schwab, M. *et al. Nature* **308**, 288-291 (1984).
16. Nepveu, A., Fahrlander, P. D., Yang, J. & Marcu, K. B. *Nature* **317**, 440-443 (1985).
17. Whitlock, C. A., Ziegler, S. F., Treiman, L. J., Stafford, J. J. & Witte, O. N. *Cell* **32**, 903-911 (1983).
18. Carney, D. N. *et al. Cancer Res.* **45**, 2913-2923 (1985).
19. Gazdar, A. F., Carney, D. N., Nau, M. F. & Minna, J. D. *Cancer Res.* **45**, 2924-2930 (1985).
20. Becker, K. L. & Gazdar, A. F. *The Endocrine Lung in Health and Disease* (Saunders, Philadelphia, 1984).
21. Nisen, P. D. *et al. EMBO J.* (submitted).
22. Gonzalez-Crussi, F. (ed.) & Franklin, W. A. in *Wilms' Tumour and Related Renal Neoplasms of Childhood*, 7 (CRC Press, Boca Raton, 1984).
23. Yancopoulos, G. D. & Alt, F. W. *Cell* **40**, 271-281 (1985).
24. Stanton, L. W., Fahrlander, P. D., Tesser, P. M. & Marcu, K. B. *Nature* **310**, 423-425 (1984).
25. Alt, F. W., Rosenberg, N., Lewis, S., Thomas, E. & Baltimore, D. *Cell* **27**, 381-390 (1981).

Human p53 gene localized to short arm of chromosome 17

Carl Miller*, T. Mohandas†, D. Wolf‡, M. Prokocimer‡, V. Rotter‡ & H. Phillip Koeffler*

* Department of Medicine, 32-139 CHS, Division of Hematology/Oncology, Factor Bldg, UCLA Medical Center, Los Angeles, California 90024, USA

† Division of Medical Genetics, Harbor-UCLA Medical Center, Torrance, California 90509, USA

‡ Department of Cell Biology, The Weizmann Institute of Science, Rehovot 76100, Israel

The p53 gene codes for a nuclear protein that has an important role in normal cellular replication¹⁻⁴. The concentration of p53 protein is frequently elevated in transformed cells^{1,2}. Transfection studies show that the p53 gene, in collaboration with the activated *ras* oncogene, can transform cells⁵⁻⁷. Chromosomal localization may provide a better understanding of the relationship of p53 to other human cellular genes and of its possible role in malignancies associated with specific chromosomal rearrangements. A recent study mapped the human p53 gene to the long arm of chromosome 17 (17q21-q22) using *in situ* chromosomal hybridization⁸. Here, by Southern filter hybridization of DNAs from human-rodent hybrids, we have localized the p53 gene to the short arm of human chromosome 17.

Cell hybrids isolated from fusions of mouse thymidine kinase (TK)-deficient cells (B82) and human fibroblasts were used to localize the p53 gene. The hybrids, which have been described previously⁹, were analysed for the presence of human p53-reactive *Eco*RI fragments. The p53 probe detected 14-kilobase (kb) fragments in *Eco*RI-cleaved human parental DNA (data not shown, but the hybridization and washing conditions are given in Fig. 1 legend). Hybridization with p53-related genes was not detected in mouse DNA when stringent washing conditions were used. Analysis of 19 mouse-human hybrid clones (data not shown) showed that the presence of human p53 sequences in these hybrids coincided only with the incidence of human chromosome 17. In two hybrid clones, human chromosome 17 was the only cytogenetically detectable human chromosome and these two clones were positive for the presence of human p53. These results assign the p53 gene to human chromosome 17, in agreement with the results of *in situ* hybridization⁸.

To map the human p53 regionally, DNAs from two Chinese hamster-human clones (17-24, 17-12), which contained the long arm but not the short arm of chromosome 17, were tested for the presence of human p53 sequences (Fig. 1). These clones were formed from the fusion of TK-deficient Chinese hamster cells and human cells containing a reciprocal translocation between chromosomes 9 and 17 (ref. 10). Hybrid clones were selected in hypoxanthine-aminopterin-thymidine (HAT) medium and found to contain either the der(17) chromosome

Received 31 October; accepted 30 December 1985.

1. Kohl, N. E. *et al. Nature* **309**, 73-77 (1986).
2. DePinho, R. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
3. Colby, W. W. *et al. Nature* **301**, 722-725 (1983).
4. Yancopoulos, G. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5455-5459 (1985).
5. Schwab, M., Varmus, H. E. & Bishop, J. M. *Nature* **316**, 160-163 (1985).
6. Nau, M. M. *et al. Nature* **318**, 69-73 (1985).
7. Varmus, H. A. *Rev. Genet.* **18**, 553-612 (1984).
8. Kohl, N. E. *et al. Cell* **35**, 359-367 (1983).
9. Schwab, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4940-4944 (1984).

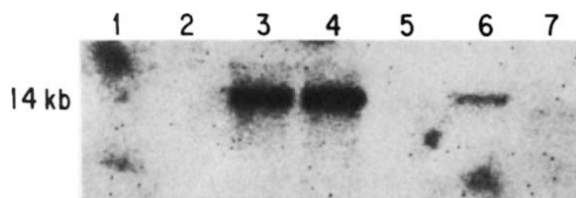


Fig. 1 Hybridization of the human p53 gene to *Eco*RI-digested DNAs from somatic cell hybrids of human and rodent. Lanes 1, 2, Chinese hamster and mouse (B82, GMO376A) parental DNAs (10 µg) respectively; lanes 3 and 4, 5 and 10 µg, respectively, of human parental DNA (IMR, Mutant Cell Repository, Camden, New Jersey); lanes 5 and 7, DNA (10 µg) from human-hamster hybrid clones 17-12 and 17-24, respectively (which do not have the short arm of human chromosome 17); lane 6, DNA (10 µg) from human-mouse hybrid clone 84-28 (whose only human chromosome is the entire chromosome 17).

Methods. After size-fractionation by electrophoresis, the *Eco*RI-digested DNAs were transferred to nylon-based paper (Biodyne; ICN, Irvine, California) and hybridized to a nick-translated human p53 cDNA clone, pp53-H13 (a 1.7-kb cDNA clone containing the 5' sequences of the message)^{14,15}. The blots were washed to high stringency with a final wash of 0.1×SSC at 65°C for 60 min. The two 14-kb p53-hybridizing fragments were determined by their positions in relation to *Hind*III-digested λ DNA fragments.

(9pter→9p11::17p11→17qter) or a normal human chromosome 17. Two independent clones (17-24, 17-12) grown in HAT medium were used for the study. Both of these clones retained the der(17) chromosome and were positive for the expression of the human enzyme markers galactose 1-phosphate uridylyl-transferase (GALT), soluble aconitase (ACO1) and adenylate kinase 3 (AK3) and negative for the expression of human adenylate kinase 1 (AK1). The loci for GALT, ACO1 and AK3 are confirmed markers on the short arm of human chromosome 9, whereas AK1 resides at the distal end of the long arm of human chromosome 9 (ref. 11). Thus the biochemical studies confirmed the cytogenetic findings, making it unlikely that the reciprocal product of the translocations consisting of the long arm of chromosome 9 and the short arm of chromosome 17 was present in these two clones. No hybrid clone was identified retaining the der(9) in the absence of either a chromosome 17 or der(17) in this series of hybrids. The hamster-human clones 17-24 and 17-12 (Fig. 1, lanes 5 and 7) were negative for human p53-hybridizing sequences, while murine-human clone 84-28 (lane 6), whose only human chromosome was the entire 17, contained p53 sequences. At the stringency used in these experiments, no hybridization with p53-related genes was detected in parental rodent DNA and only the two identical p53-hybridizing fragments were observed in human DNA. Taken together, these results show that the human p53 gene lies on the short arm of human chromosome 17. The reason for the discrepancy between our results and those of Le Beau *et al.*⁸ is not clear but may relate to their use of a murine p53 complementary DNA clone for their *in situ* chromosome studies; the murine and human p53 genes are not completely homologous.

Knowledge of the locations of p53 and other oncogene-related genes is clearly important for determining the relationship between known chromosomal aberrations in tumours and oncogenes found on these chromosomes. Patients with blast crisis of chronic myelogenous leukaemia, preleukaemia and acute myelogenous leukaemias frequently have one of several abnormalities involving the short arm of chromosome 17, especially an isochromosome of the long arm of chromosome 17 [i(17q)] which results in monosomy of the short arm of chromosome 17 (ref. 12). The relationship between leukaemia and p53 expression remains to be determined.

After submission of this manuscript, Benchimol *et al.* also

reported localization of p53 to the short arm of human chromosome 17 (ref. 13).

This work was supported in part by NIH grants CA 26038, CA 32737 and CA 33936, the Dr Murray Geisler Memorial Fund and the Louis Fagin Leukaemia Research Foundation (to H.P.K.). H.P.K. has a Career Development Award from the NIH and is a member of the Jonsson Cancer Center. C.M. is a fellow of the Bank of America-Giannini Foundation. D.W., M.P. and V.R. were supported by a grant from the Cancer Research Institute and the Weizmann Institute of Science. V.R. has the Norman and Helen Asher Career Development Chair. We thank Regina Simon and Suzanne Bookstaver for secretarial help; and John Ranyard and George Helfand for technical skills.

Received 12 September; accepted 6 December 1985.

1. Crawford, L. V. *Int. Rev. exp. Path.* **25**, 1-50 (1983).
2. Rotter, V. & Wolf, D. *Adv. Cancer Res.* **43**, 113-140 (1985).
3. Milner, J. & Milner, S. *Virology* **112**, 785-788 (1981).
4. Reich, N. C. & Levine, A. J. *Nature* **308**, 199-201 (1984).
5. Parada, L. F., Land, H. A., Weinberg, R. A., Wolf, D. & Rotter, V. *Nature* **312**, 649-651 (1984).
6. Eliyahu, D., Roz, A., Gruss, P., Givol, D. & Oren, M. *Nature* **312**, 646-649 (1984).
7. Jenkins, J., Rudge, K. & Currie, G. *Nature* **312**, 651-653 (1984).
8. Le Beau, M. M., Westbrook, C. A., Diaz, M. O., Rowley, J. D. & Oren, M. *Nature* **316**, 826-828 (1985).
9. Taggart, R. T., Mohandas, T., Shows, T. B. & Bell, G. I. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6240-6244 (1985).
10. Mohandas, T. *et al. Am. J. hum. Genet.* **31**, 586-600 (1979).
11. Meerakhan, P. & Smith, M. *Cytogenet. Cell Genet.* **37**, 71-102 (1984).
12. Mitelman, F. *Progress and Topics in Cytogenetics* Vol. 5 (ed. Sandberg, A. A.) 1-722 (Liss, New York, 1985).
13. Benchimol, S. *et al. Somat. Cell molec. Genet.* **11**, 505-509 (1985).
14. Wolf, D., Laver-Rudich, Z. & Rotter, V. *Molec. cell. Biol.* **5**, 1887-1893 (1985).
15. Wahl, G., Stern, M. & Stark, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3688 (1979).

Differential RNA splicing predicts two distinct nerve growth factor precursors

R. H. Edwards*†, M. J. Selby* & W. J. Rutter*‡

* Hormone Research Institute and Department of Biochemistry & Biophysics, and † Department of Neurology, University of California, San Francisco, California 94143, USA

Nerve growth factor (NGF) has a crucial role in the development of sensory and sympathetic neurones¹⁻⁶. However, although it can affect other neural cell types under certain experimental conditions⁷, no biological role has been convincingly demonstrated elsewhere in the nervous system. The 5' end of the mouse NGF gene contains several relatively short exons⁸. The NGF messenger RNA contains two in-frame initiator methionine codons; the second precedes the signal peptide sequence. Studies of the translation of other eukaryotic mRNAs indicate that the first AUG is preferred, suggesting that the signal for secretion might be ambiguous. We have analysed the NGF mRNA species from various cell types, some of which (clonal myoblast and fibroblast cell lines) are known to secrete NGF^{9,10}, to search for different NGF transcripts. One pathway of RNA splicing generates the transcript already described from a submaxillary gland complementary DNA clone¹¹. We demonstrate here that there is another splicing pathway, leading to a shorter transcript that lacks the second exon. This short transcript is the major form in most other mouse tissues and in the tissues of several other species, but both transcripts are usually present. In the short transcript, the initiator methionine is immediately upstream from a signal peptide-like sequence whereas in the long transcript the first methionine is 62 amino acids upstream from the signal peptide-like sequences. This may result in a different cellular localization of the NGF or alter the biological activity of the NGF precursor.

‡ To whom correspondence should be addressed.

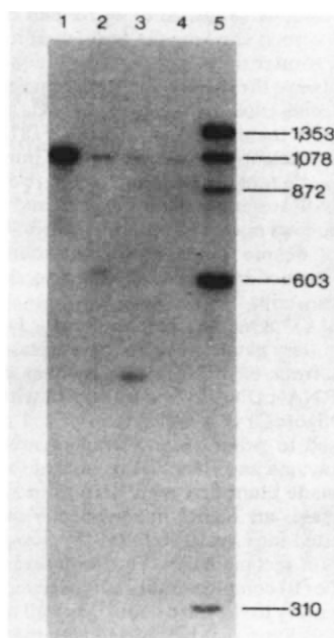


Fig. 1 S_1 nuclease analysis of the 5' end of NGF mRNA. Lane 1, undigested probe; lane 2, 20 μ g female mouse submaxillary gland total RNA; lane 3, 20 μ g poly(A)⁺ G8 cell RNA; lane 4, 20 μ g yeast transfer RNA; lane 5, M_r standards.

Methods. RNA was prepared in guanidinium isothiocyanate²⁸, precipitated in lithium chloride²⁹, extracted and reprecipitated in ethanol. Poly(A)⁺ RNA was selected on oligo(dT)-cellulose³⁰. A NGF cDNA clone in pBR328 was digested with *Nco*I (+616), phosphatased and 5' end-labelled with [γ -³²P]ATP and polynucleotide kinase. The DNA was recut with *Hind*III and the 5' NGF fragment isolated on an 8% acrylamide gel. The probe (5,000 c.p.m.) was heated with the RNA in 80% formamide at 85 °C for 15 min, then hybridized at 45 °C for 3 h. S_1 nuclease digestion with 5 U of enzyme (PL Biochemicals) was performed at 37 °C for 1 h. The samples were extracted, precipitated twice and loaded onto a 5% denaturing acrylamide gel³¹.

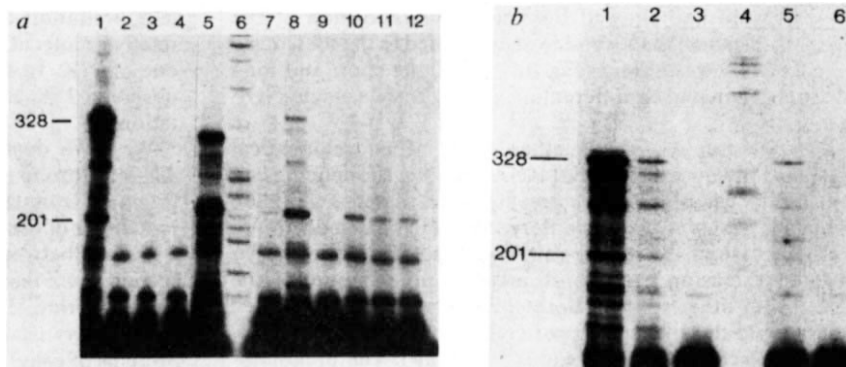
In the early experiments we found that the mouse G8 embryonic myoblast cell line contains NGF mRNA by Northern blotting. Using the cloned mouse submaxillary gland NGF cDNA¹¹ as a probe, we observed after high stringency washing a band representing about 0.001% of the cell mRNA and migrating at about 1,300 bases, approximately the size of NGF transcripts¹². However, the 5' end of the NGF gene contains several

small (32–127-nucleotide) exons (M.J.S., in preparation), thus alternative splicing might yield mRNAs of slightly differing size that are not resolved on these gels. We therefore tested, using S_1 nuclease digestion, whether NGF mRNAs contained divergent 5' ends. A 5' end-labelled fragment extending from the *Nco*I site at +616 in the cDNA into the plasmid sequences flanking the insert was hybridized to poly(A)⁺ RNA from G8 myoblast cells. After digestion with S_1 nuclease, several protected fragments were detected (Fig. 1). The longest protected fragment corresponds to the original submaxillary gland cDNA clone¹¹; the shortest fragment, greatest in intensity, corresponds to a fragment cleaved very close to the splice junction at position +158. We observe a similar set of protected fragments in experiments using total RNA from female mouse submaxillary gland, but the longest fragment predominates.

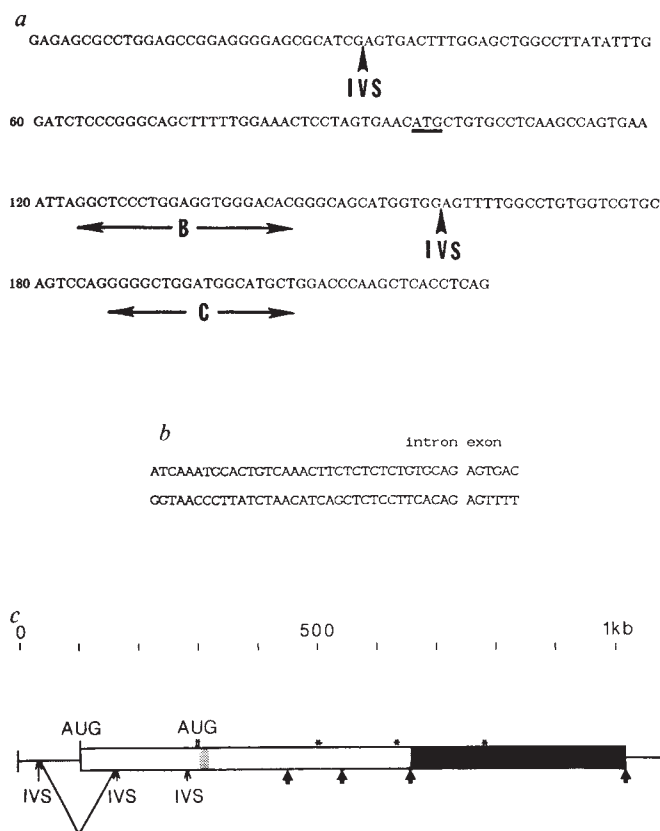
Primer extension experiments verified the presence of the short and long transcripts and several minor extension products in several tissues. In G8 myoblasts, the strongest band corresponds to a transcript which initiates about 35 bases 5' to the splice junction at +158 (Fig. 2a, lane 8), thus being shorter than and different from the original cDNA. In submaxillary gland RNA the most intense band is the longest and corresponds to the size expected for the original cDNA clone (Fig. 2a, lane 1). Similar experiments demonstrate the presence of short and long transcripts in various tissues from mouse, rat, guinea pig and human. The short NGF transcript is the predominant one in mouse L cells, placenta and in hindlimb at several stages of early development (Fig. 2a, lanes 8, 3, 10–12), in rat submaxillary gland, placenta (Fig. 2a, lanes 2, 4) and developing brain (Fig. 2b, lane 6). However, the short and long transcripts occur in about equal concentrations in human placenta (Fig. 2b, lane 5). The additional bands present especially in the submaxillary RNA probably represent stops to reverse transcription. The finding of relatively high levels of NGF mRNA in placenta is unexpected. Shelton and Reichardt¹² have shown a correlation between the levels of NGF mRNA and the extent of sympathetic innervation. The placenta is an exception to this rule; it has relatively high levels of NGF but is not innervated. Guinea pig prostate, another major source of NGF¹³, contains the short transcript, but also displays an intense band somewhat shorter than the long transcript observed in the mouse salivary gland (Fig. 2a, lane 5). This finding is of interest because mature NGF purified from guinea pig prostate also displays different antigenic properties from mouse NGF¹³.

To establish whether the shortest extension product is a true NGF transcript or the result of premature termination, we sequenced the G8 myoblast RNA primer extension product. The sequence was identical with the known NGF cDNA up to

Fig. 2 **a**, Primer extension assays of 20 μ g female mouse submaxillary gland total RNA (lane 1), 20 μ g female rat submaxillary gland total RNA (lane 2), 20 μ g mouse placental total RNA (lane 3), 20 μ g rat placental total RNA (lane 4), 20 μ g guinea pig prostate total RNA (lane 5), *Hpa*II digest of pBR322 (lane 6), 20 μ g fused G8 cell total RNA (lane 7), 1 μ g L-cell poly(A)⁺ RNA with 19 μ g tRNA (lane 8), 1 μ g poly(A)⁺ RNA from NIE-115 cells (which lack NGF transcripts by Northern blotting) with 19 μ g tRNA (lane 9), and 20 μ g total RNA from mouse hind limb at 17 days of gestation, birth and 2 days later (lanes 10–12). **b**, Primer extension assays using 10 μ g total RNA from male (lane 1) and female (lane 2) mouse submaxillary glands, 20 μ g tRNA (lane 3), 5 μ g poly(A)⁺ RNA from human placenta (lane 5) and 15-day rat brain (lane 6); standards in lane 4 are a *Hae*III digestion of ϕ X174 DNA.



Methods. RNA was prepared as for Fig. 1. The primer was prepared by annealing a 22-nucleotide primer (complementary to +328 to +306 in the original cDNA) to an M13 template containing NGF, extending with the Klenow fragment of DNA polymerase I in the presence of all four α -labelled deoxyribonucleotides, digesting with *Sph*I (which cleaves at position +199) and separating the strands on a 5% acrylamide-urea gel³². The 129-base, single-stranded primer (1.5×10^5 c.p.m. having a specific activity $\sim 10^9$ c.p.m. per μ g) was then annealed to the RNA and extended with reverse transcriptase according to the procedure of McKnight and Kingsbury.



the +158 splice junction; it was followed by an interfering, unreadable ladder of ~35 bases (data not shown), then a faint signal corresponding to the known NGF sequence resumed. This pattern could have resulted from two mRNAs differing in sequence 5' to the +158 splice site.

We sought alternative NGF transcripts by screening a mouse submaxillary gland cDNA library specifically constructed to conserve the 5'-terminal sequences (see Fig. 3 legend). (Although this tissue has a lower proportion of the short transcript, it contains more of both forms than any other tissue known.) The unamplified library (10^6 plaque-forming units in λ gt10) was screened with the NGF cDNA under high stringency, and with oligonucleotide probes from the second and third exons, respectively (Fig. 3a). Of 1,000 clones that hybridized with the cDNA, 100 hybridized with the oligonucleotide probe from the third exon but not with the probe from the second. Three such clones were sequenced and showed the precise juxtaposition of the first and third exons; the second exon contained in the previously reported cDNA was absent (Fig. 3c). Thus, the short and long transcripts originated by differential splicing from a single NGF gene.

The expression of several other genes of neurobiological interest also involves differential RNA splicing; the splicing can be extremely complex, developmentally regulated and tissue specific¹⁴⁻¹⁸. We infer that the differential splicing demonstrated here in the case of the mouse salivary gland also accounts for the primer extension patterns observed in other tissues of the mouse and in other species. Both splicing pathways described here adhere to the splice acceptor consensus sequence at the 5' ends of the second and third exons¹⁹ (Fig. 3b). The predominance of the long mRNA in the male mouse submaxillary gland (from which the first NGF cDNA clones were derived) obviously decreased the probability of isolating the short NGF cDNA; in addition, the primer used by Scott *et al.*¹¹ to sequence the 5' end of the transcript adventitiously straddled the +158 splice junction and so precluded the detection of the alternative message. The mechanism involved in this differential splicing

Fig. 3 a, Nucleotide sequence of the 5' end of the larger NGF transcript with arrows showing the position of intervening sequences (IVS). The shorter transcript described in the text lacks all the nucleotides between these IVSs. **b**, An oligonucleotide complementary to the second exon (+145 to +124); **c**, an oligonucleotide complementary to the third exon (+205 to +187)³⁴, used to screen the library. **b**, Nucleotide sequence of the introns just 5' to the second and third exons. **c**, Structure of the two NGF transcripts. The thin line indicates untranslated regions and the box translated regions. Thin arrows mark IVS and thick arrows potential cleavage sites; asterisks denote potential glycosylation sites. Length is measured in nucleotides (above), with zero the cap site of the larger RNA transcript.

Methods. Poly(A)⁺ RNA was prepared as for Fig. 1 from the male mouse submaxillary gland. Reverse transcriptase synthesized first-strand cDNA from the RNA using random calf thymus DNA primers. The RNA-cDNA hybrid was tailed with dG and terminal transferase. Oligo(dC) in a molar ratio of 2:1 relative to the first strand was used to prime second-strand synthesis following the method of Okayama and Berg³⁵. The ends of the double-stranded cDNA were made blunt first with Klenow enzyme, then with T4 DNA polymerase; an *Eco*RI linker adapter was added and the sequence inserted into λ gt10 (ref. 36). *In vitro* packaging yielded a library of $\sim 10^6$ recombinants. The library was screened with an oligonucleotide (B) complementary to the second exon and another (C) complementary to the third exon³⁴, as well as the NGF cDNA. Of almost 1,000 clones that hybridized with the cDNA under high stringency and with oligonucleotides B and C, only about 100 hybridized with C but not B. Six of these were subcloned into M13 and the nucleotide sequence was determined by the dideoxy method in both directions³⁷. Three showed a precise deletion of the second exon, with otherwise complete agreement to the original sequences, down to the 5' end of the mRNA. Other clones showed premature termination of the larger transcript between oligonucleotides C and B.

may result from different relative concentrations of *trans*-acting elements present in ribonucleoprotein particles that interact with these two splice acceptors^{20,21}. Biological significance for the two pathways therefore demands a demonstration of different function for the two transcripts.

The absence of the second NGF exon may affect the biological properties of the NGF precursor. It produces a 5'-untranslated region of 172 bases and removes the first AUG. The initiation of translation would presumably occur at the next in-frame AUG. Both of these potential initiation codons have reasonable consensus sequences according to the Kozak rules²². Interestingly, in the short transcript, an out-of-frame AUG with good consensus sequence appears first in the mRNA, but a stop codon in the same frame immediately precedes the in-frame AUG. Translation of the short and long transcripts may explain more fully the observed complex pattern of the NGF precursor protein. Initiation at the first AUG would result in a protein of relative molecular mass (M_r) 34,000 (34K) and at the second, one of 27K. In salivary gland slices, Darling and Shooter have observed 35K and 29K NGF precursors after immunoprecipitation.

We have demonstrated use of the two potential initiation AUG codons to produce two distinct NGF proteins. A restriction fragment from the NGF cDNA containing both AUGs and the remainder of the protein coding sequence, and another fragment truncated between the two AUGs and thus containing only the second, were inserted into the SP65 vector system. Transcription of RNA from these templates and comparison of the protein products translated *in vitro* shows that the longer RNA yields proteins of roughly 34K and 27K M_r (corresponding to initiation of transcription at the first and second AUGs, respectively) (Fig. 4, lanes 2, 3) whereas the truncated form gives only the 27K protein. Truncation of both NGF RNAs at their 3' end (by linearization of the template DNA with *Nco* and *Sca* at +616 and +810, respectively, in the cDNA) shortens the translated proteins by the expected amount, further establishing the origins of translation (Fig. 4, lanes 4-7). Although these *in vitro* results

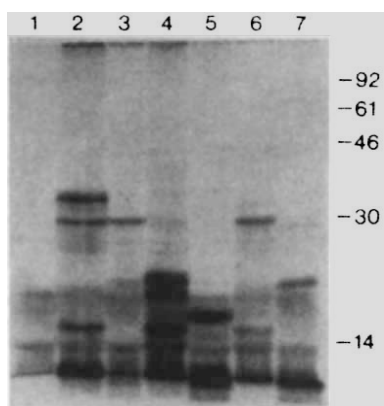


Fig. 4 *In vitro* translation of NGF RNA. For the longer construction containing both potential initiator AUGs (at +98 and +296 in the nucleotide sequence of the NGF mRNA), the cDNA was digested with *Sma*I (+66) and *Pst*I (+1027) and inserted into the SP65 vector. For the shorter construction containing only the second potential initiation codon, the cDNA was partially digested with *Bst*XI (+184), then *Pst*I and inserted into the same vector. Following linearization with the enzymes designated below, the two constructions (named LmRNA and SmRNA for the longer and shorter, respectively) were then used as templates for RNA synthesis with the SP6 polymerase³⁸. This RNA was then translated *in vitro* by the wheat germ system in the presence of ³⁵S-methionine³⁹. The protein products were precipitated in 10% trichloroacetic acid and subjected to electrophoresis on 12% SDS-polyacrylamide. The gel was enhanced with 1 M sodium salicylate and exposed overnight at -70 °C with an enhancing screen, Lane 1, no RNA; lane 2, LmRNA linearized with *Hind*III (3' to the inserted cDNA); lane 3, SmRNA linearized with *Hind*III; lane 4, LmRNA linearized with *Nco*I (at +616 in the cDNA); lane 5, SmRNA linearized with *Nco*I; lane 6, LmRNA linearized with *Sca*I (at +810 in the cDNA); lane 7, SmRNA linearized with *Sca*I; numbers on the right indicate the M_r s of the protein standards ($\times 10^3$).

do not prove alternative use of the two AUGs *in vivo*, they do show that the translational apparatus can recognize both AUGs.

An understanding of the biological significance of differential splicing in the expression of NGF requires further information about the functions of the various precursor molecules. The NGF precursor contains several di- and tetrabasic residues that are potential peptidase cleavage sites (Fig. 3c). The small NGF precursor protein derived from the shorter transcript retains these sites, but the potential NH₂-terminal cleavage product would be truncated. If the full precursor or one of its derived peptides has biological activity, the two RNA splicing pathways could result in differing biological effects. It is intriguing that the two transcripts result in proteins with the hydrophobic sequence in different positions; in the long precursor, it resides in the middle of the molecule and might serve as a membrane anchor (similar to the epidermal growth factor precursor in the kidney²⁴); in the short precursor, the hydrophobic region is at the N-terminus, and may act as a signal peptide to ensure secretion²⁵. Differential splicing could therefore affect cellular localization, as with IgM and yeast invertase^{26,27}. Although the short transcript presumably results in a secreted NGF that can act at a distance, the long transcript might produce a form of NGF that is membrane-bound and could interact with high efficiency at a receptor on a contiguous cell; these mechanisms could provide the basis for direct cell-cell communication.

We thank Dr David Standing for advice on the RNA analysis, Drs Jang Han and Jeff Edman for help with cDNA synthesis, Drs James Ou, Michael Walker and Eric Fodor for their thoughtful discussion, Drs Pablo Garcia and Peter Walter for assistance with *in vitro* translation, Dr Philip Barr for the oligonucleotides, and Ms Leslie Spector for preparing the manuscript. R.H.E.

was supported by a postdoctoral fellowship from the NINCDS (NS07529-02). This research was supported by NIH grant AM 21334 (to W.J.R.).

Received 20 August 1985; accepted 2 January 1986.

- Levi-Montalcini, R. & Angeletti, P. U. *Pharmac. Rev.* **18**, 619-628 (1966).
- Gorin, P. D. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5382-5386 (1979).
- Stockel, K., Schwab, M. E. & Thoenen, H. *Brain Res.* **99**, 1-16 (1975).
- Banks, B. E. C. *et al. J. Physiol., Lond.* **247**, 289-298 (1975).
- Chun, L. L. Y. & Patterson, P. H. *J. Cell Biol.* **75**, 694-704 (1975).
- Thoenen, H. & Barde, Y.-A. *Physiol. Rev.* **60**, 1284-1335 (1980).
- Mobley, W. C. *et al. Science* **229**, 284-287 (1985).
- Ullrich, A. *et al. Nature* **303**, 821-825 (1983).
- Murphy, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 4496-4500 (1977).
- Oger, J. O. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 1554-1558 (1974).
- Scott, J. *et al. Nature* **302**, 538-540 (1983).
- Shelton, D. L. & Reichardt, L. F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7951-7955 (1984).
- Harper, G. P. *et al. Nature* **279**, 160-162 (1979).
- Herbert, E. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 375-384 (1983).
- Amara, S. G. *et al. Nature* **298**, 240-244 (1982).
- Nawa, H., Kotani, H. & Nakanishi, S. *Nature* **312**, 729-734 (1984).
- Falkenthal, S., Parker, V. P. & Davidson, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 449-453 (1985).
- Medford, R. M. *et al. Cell* **38**, 409-421 (1984).
- Mount, S. *Nucleic Acids Res.* **10**, 459-472 (1982).
- Steitz, J. A. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 893-900 (1983).
- Mount, S. M. *et al. Cell* **33**, 509-518 (1983).
- Kozak, M. *Nucleic Acids Res.* **9**, 5233-5252 (1981).
- Darling, T. L. J. *et al. Cold Spring Harb. Symp. quant. Biol.* **48**, 427-534 (1983).
- Rall, L. *et al. Nature* **313**, 228-233 (1985).
- Lingappa, V. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2338-2342 (1978).
- Early, P. *et al. Cell* **20**, 313-319 (1980).
- Carlson, M. & Botstein, D. *Cell* **28**, 145-154 (1982).
- Chirgwin, J. M. *et al. Biochemistry* **18**, 5294-5297 (1979).
- Cathala, G. *et al. DNA* **2**, 329-335 (1983).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721-732 (1977).
- Hu, N. T. & Messing, J. *Gene* **17**, 271-277 (1982).
- McKnight, S. L. & Kingsbury, R. *Science* **217**, 316-324 (1982).
- Wallace, R. B. *et al. Nucleic Acids Res.* **9**, 2647-2656 (1981).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161-170 (1982).
- Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D.) pp. 49-78 (IRL Press, Oxford, 1984).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057-7070 (1984).
- Erickson, A. H. & Blobel, G. *Meth. Enzym.* **96**, 38-50 (1983).

Cell contact- and shape-dependent regulation of vinculin synthesis in cultured fibroblasts

Florette Ungar*, Benjamin Geiger† & Avri Ben-Ze'ev*

Departments of *Genetics and †Chemical Immunology, Weizmann Institute of Science, Rehovot, Israel 76100

Recent studies have demonstrated the fundamental role of cell-substrate contacts and changes in cell shape in the regulation of cell growth, motility and differentiation¹⁻⁷, but the molecular basis for these phenomena is poorly understood. Because of the involvement of cytoskeletal networks in cell morphogenesis and contact formation, it is of interest that the expression of genes encoding several cytoskeletal proteins is markedly affected by changes in cell contacts and configuration⁶⁻¹⁰. Because most of these phenomena involve changes in the form, extent or topology of cell contacts, we sought to determine whether the expression of components directly involved in the formation of cell-cell or cell-substrate contacts is affected by the respective cellular interactions. A suitable candidate for such analysis is vinculin, a cytoskeletal protein of relative molecular mass (M_r) 130,000 (130K), which is localized in focal contacts¹¹⁻¹³ and intercellular adherens junctions¹⁴. The assembly of vinculin into a membrane-bound junctional plaque seems to be one of the earliest cellular responses to contact with exogenous substrates, leading to the subsequent local assembly of the actin-rich microfilament bundles^{15,16}. Here we report on the regulation of vinculin synthesis in response to environmental conditions that affect cell shape and contacts.

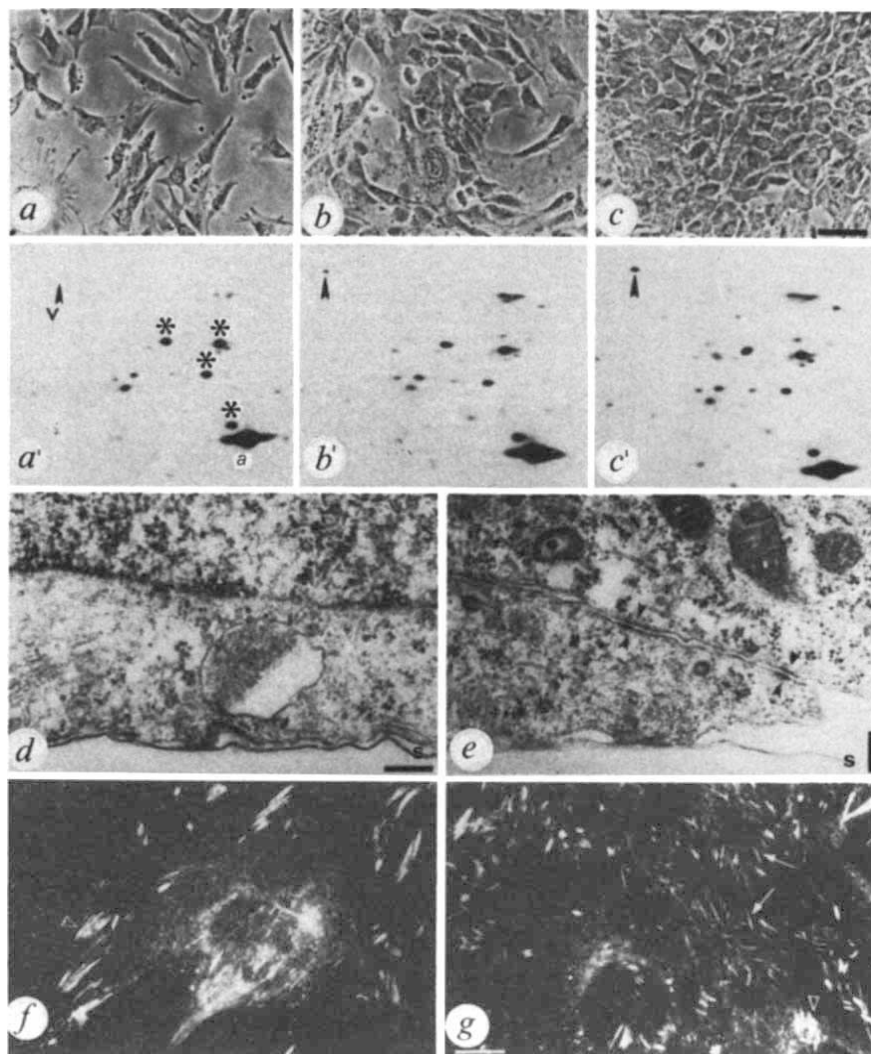


Fig. 1 Vinculin synthesis and cell-contact formation in 3T3 cells plated at different densities. *a-c*, Phase-contrast photomicrographs of Swiss 3T3 fibroblasts plated for 2 days at concentrations of 2×10^4 (*a*), 10^5 (*b*) and 5×10^5 (*c*) cells per 35-mm Falcon tissue-culture dish. *a'-c'*, Analysis of proteins from corresponding samples labelled for 4 h with $100 \mu\text{Ci ml}^{-1}$ ^{35}S -methionine. Extracts were prepared by scraping cells into a buffer containing 1.0% Triton X-100, 0.6 M KCl, 5 mM EGTA, 7 mM 2-mercaptoethanol, 1 mM phenylmethylsulphonyl fluoride and 10 mM Tris-HCl (pH 7.4) and combining the supernatant with that obtained after treatment of the pellet with DNase I and then with 0.6 M KCl. Equal amounts of TCA-insoluble radioactivity (5×10^5 c.p.m. per sample) were analysed by two-dimensional gel electrophoresis as described previously²⁵. *a*, Actin; *v* and arrowheads point to vinculin. Asterisks mark proteins whose intensity of labelling did not change and which served as reference spots. *d, e*, Electron micrographs of 3T3 cells plated at a density of either 2×10^4 cells per dish (*d*) or 5×10^5 cells per dish (*e*) for 2 days. Cells were fixed in the dish with 1% glutaraldehyde in 0.1 M cacodylate buffer pH 7.2, postfixed in 1% OsO_4 , flat-embedded in Epon and sectioned perpendicular to the plane of the dish. *s*, Substrate; arrowheads point to intercellular adherens-type junctions with submembrane electron density. *f, g*, Indirect immunofluorescence staining of sparse (*f*) and dense (*g*) 3T3 cultures with a monoclonal antibody against chicken gizzard vinculin and rhodaminated goat anti-mouse IgG. Arrowheads point to large vinculin-containing adhesion plaques; arrows mark thin vinculin-containing structures corresponding to cell-cell contacts. Scale bars: *a-c*, 50 μm ; *d, e*, 0.2 μm ; *f, g*, 10 μm .

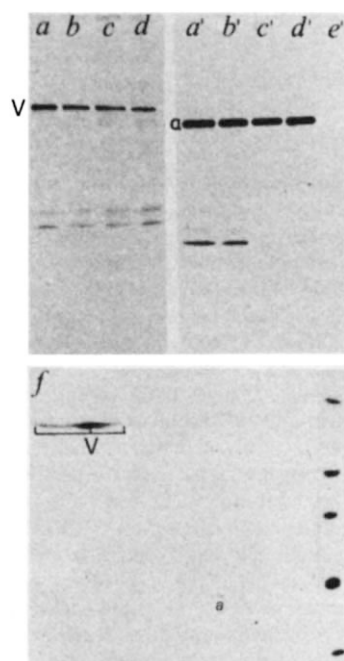
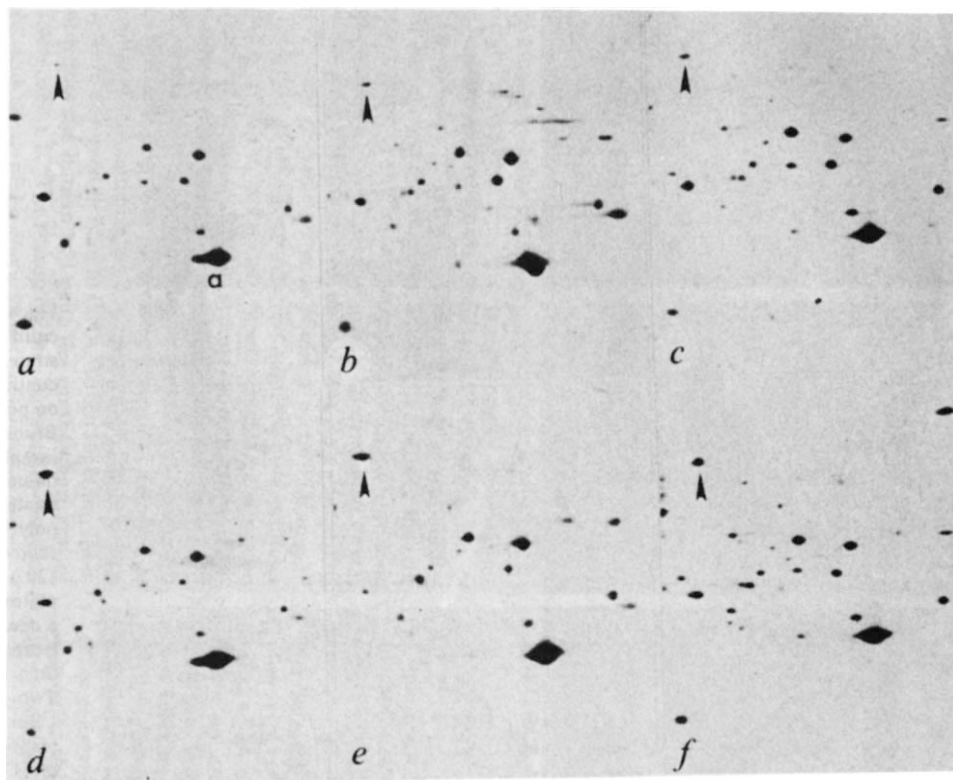


Fig. 2 Identification of vinculin in sparse and dense cultures of 3T3 cells by immunoprecipitation with a monoclonal, vinculin-specific antibody. 3T3 cultures on 35-mm dishes which had been seeded with either 5×10^5 cells (*a, a', e'*), 2.5×10^5 cells (*b, b'*), 10^5 cells (*c, c'*) or 2×10^4 cells (*d, d'*) were labelled for 4 h with ^{35}S -methionine and lysed in RIPA buffer (10 mM NaH_2PO_4 , 100 mM NaCl, 1.0% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, pH 7.0). We took 2×10^6 c.p.m. from each culture for immunoprecipitation with either anti-vinculin antibody (*a-d, f*), anti- α -actinin antibody (*a'-d'*) or non-immune mouse immunoglobulin (*e'*). The immune complexes were precipitated by binding to *Staphylococcus aureus*, according to Kessler²⁶. *a-d, a'-e'*, Analysis on 10% acrylamide gels; *f*, two-dimensional gel analysis of immunoprecipitate from 5×10^5 cells labelled as above. α , α -Actinin; *V*, vinculin (all isoforms); *a*, actin. Relative molecular mass markers for the gel include (top to bottom) myosin heavy chains, phosphorylase *b*, bovine serum albumin, ovalbumin and carbonic anhydrase.

Fig. 3 Vinculin synthesis in sparse and dense 3T3 cultures determined by ^{35}S -methionine pulses and by pulse-chase experiments. 3T3 cells plated at 2×10^4 cells per 35-mm dish (*a-c*) or at 5×10^5 cells per 35-mm dish (*d-f*) were either labelled for 45 min with $400 \mu\text{Ci ml}^{-1}$ of ^{35}S -methionine (*a, d*) or labelled for 45 min as above and then washed extensively and further incubated in medium containing unlabelled methionine for 10 h (*b, e*). Alternatively, cells were continuously labelled for 11 h with $40 \mu\text{Ci ml}^{-1}$ of ^{35}S -methionine (*c, f*). Cell extracts containing 5×10^5 c.p.m. from each sample were analysed by two-dimensional gel electrophoresis as in Fig. 1. Arrowheads point to vinculin; reference spots were as in Fig. 1; *a*, actin.



To examine the role of specific cell contacts in the regulation of vinculin expression, we plated primary chicken fibroblasts or Swiss 3T3 cells at increasing densities on tissue-culture dishes. At a low plating density, cells were spread out on the substrate, rarely contacting each other. At higher densities, cell-cell contacts became prominent, the extent of spreading was limited and most cells assumed an elongated, cylindrical shape (Fig. 1 *a-c*). To compare the rate of vinculin synthesis by the different cultures, cells were labelled with ^{35}S -methionine, equal trichloroacetic acid (TCA)-precipitable labelled polypeptides were resolved by two-dimensional gel electrophoresis and the radioactivity associated with the vinculin spots was measured and normalized against several standard polypeptides on the same gels (Fig. 1*a'-c'*). Vinculin labelling was enhanced sevenfold when cell density was increased (Table 1).

Electron microscope examination of sparse and dense cultures of 3T3 cells revealed that the ventral membrane of sparse cells (Fig. 1*d*) was usually closer to the substrate than that of cells in dense cultures (Fig. 1*e*) (mean distance from the substrate = 25 nm, compared with ~60 nm in the latter). In dense cultures, however, many focal contacts were detected, as well as areas of tight cell-cell contact (arrowheads in Fig. 1*e*). Immunofluorescence staining with anti-vinculin antibody revealed that the protein was associated mainly with large ventral focal contacts in sparse cultures of 3T3 cells, and in dense monolayer cultures vinculin staining was observed both in small focal contacts and in many cell-cell contacts located above the plane of the substrate (Fig. 1*g*, arrows), in agreement with the electron microscope observations. To substantiate the suggestion that cell density affects vinculin synthesis, cell extracts of ^{35}S -methionine-labelled cultures at different densities were subjected to two-dimensional gel analysis or to immunoprecipitation with either anti-vinculin or anti- α -actinin (used as a control). The results confirmed that the 130K protein, whose labelling was augmented by the increase in cell density, was immunochemically (Fig. 2*a-d*) and biochemically (Fig. 2*f*) identical to vinculin. In chicken fibroblasts, which displayed an essentially identical

increase in vinculin expression (Table 1), we also examined the iodinated peptide map of the 130K spot and found it to be identical to that of purified chicken gizzard vinculin (data not shown).

To determine whether the changes in incorporation of ^{35}S -methionine into vinculin are caused by altered rates of synthesis or of turnover, we labelled sparse and dense 3T3 cultures for various periods of time and also carried out pulse-chase experiments (Fig. 3). Densitometric examination of the vinculin spots in comparison with control proteins (see Fig. 1) indicated that the ratio between vinculin labelling in the dense and sparse

Table 1 Relative rates of vinculin synthesis as a function of cell density and cell spreading

Cell type	Plating density per 35 mm dish	Relative rate of vinculin synthesis	Poly(HEMA) concentration ($\mu\text{g ml}^{-1}$)	Relative rate of vinculin synthesis
3T3	2.0×10^4	1.0	None	1.0
	2.0×10^5	3.75	72	0.53
	3.0×10^5	4.5	96	0.63
	5.0×10^5	7.0	120	0.28
			1,200	0.21
Chicken gizzard	2.5×10^4	1.0	ND	ND
	10^5	2.7	ND	ND
	2.5×10^5	5.3	ND	ND
	10^6	6.8	ND	ND

Cells were plated on 35-mm diameter dishes and after 2 days were labelled with ^{35}S -methionine as described in Fig. 1 legend. The intensity of the vinculin spot on two-dimensional gels was determined by densitometric scanning and normalized against several reference spots whose intensity did not change (asterisks, Fig. 1). 3T3 cells were also seeded on poly(HEMA)-coated plates as outlined in Fig. 4, labelled with ^{35}S -methionine and analysed on two-dimensional gels. ND, not done.

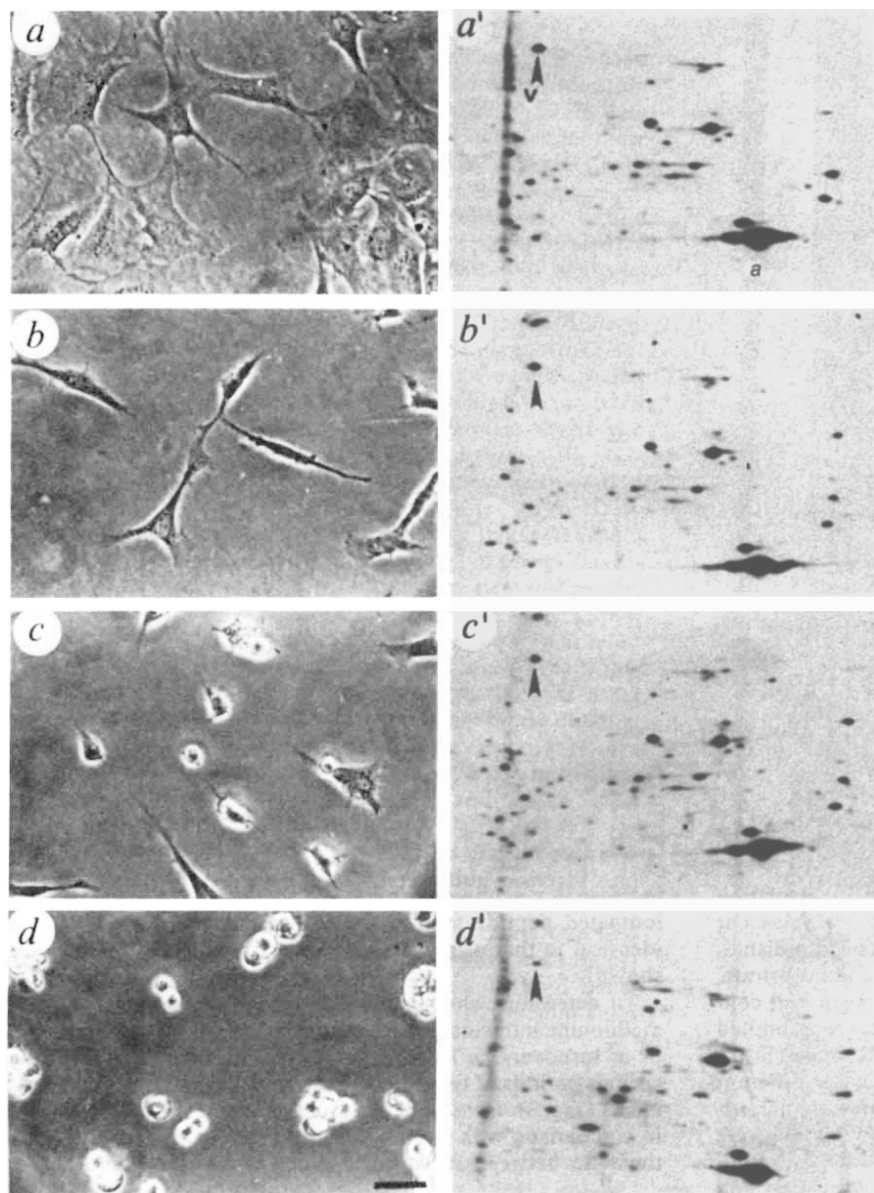


Fig. 4 Analysis of cell morphology and vinculin synthesis in 3T3 cells plated on substrates of varying adhesiveness. *a-d*, Phase-contrast photomicrographs of 3T3 cells grown on poly(HEMA) (Hydron Laboratories, New Brunswick) coated plastic plates prepared essentially as described elsewhere^{2,3}. Plastic tissue-culture plates (5 cm diameter) were treated with 2.5 ml of either ethanol (*a*) or poly(HEMA)-ethanolic solutions of the following concentrations: 96 $\mu\text{g ml}^{-1}$ (*b*); 120 $\mu\text{g ml}^{-1}$ (*c*); or 1,200 $\mu\text{g ml}^{-1}$ (*d*). The plates were dried at 37 °C for 2 days. Cells at a density of 10^5 per plate were seeded on the treated substrates and incubated for 2 days, then labelled as described for Fig. 1. *a'-d'*, Two-dimensional gel electrophoresis of the ³⁵S-methionine-labelled proteins from the respective samples, containing equal amounts of TCA-insoluble radioactivity. Cell labelling and cell extracts were prepared as described in Fig. 1. Arrowheads point to vinculin; *a*, actin; reference proteins, whose intensity of labelling did not change, are as in Fig. 1.

cultures was not significantly affected by the duration of the labelling period (Fig. 3*a, b, d, e*) or in pulse-chase experiments (Fig. 3*c, f*). We therefore conclude that the extent of cell-cell interactions affects the synthesis rather than the turnover of vinculin.

To determine whether the rate of vinculin synthesis is affected by changing the extent of cell-substrate contact, we plated equal numbers of 3T3 cells on substrates of decreasing degrees of adhesiveness, using dishes coated with different concentrations of poly(2-hydroxyethyl methacrylate) (poly(HEMA); Fig. 4*a-d*; also refs 2, 3). Two days after plating, cell shape varied from well-spread on thin poly(HEMA) films (Fig. 4*a*) to spherical on dense poly(HEMA) films (Fig. 4*d*). The level of vinculin synthesis per constant amount of total ³⁵S-labelled protein was approximately fivefold lower in spherical, non-adherent cells (Fig. 4*d'*) than in flat monolayer cells (Fig. 4*a'*), with intermediate values in cells showing intermediate degrees of spreading (Fig. 4*b', c'*; Table 1). Control experiments similar to those of Fig. 3 verified that vinculin synthesis rather than its turnover, was affected, suggesting that vinculin synthesis corre-

lates with the degree of cell spreading induced by alterations of substrate adhesiveness.

The experiments described here indicate that cells express vinculin in relation to its pattern of organization in the cell and/or to the mode of cell contacts formed. The organization of vinculin is altered in dense cultures, and the protein is extensively associated with areas of cell-cell contact of the adherens type. Thus, vinculin synthesis, like that of several other cytoskeletal proteins (vimentin, actin, tubulin and cytokeratin⁸⁻¹¹), is sensitive to the degree of cell-cell and/or cell-substrate contact (see refs 6, 17 for reviews). We suggest that, as well as its unique regulation by intercellular contacts, vinculin has, in common with the cytoskeletal proteins mentioned above, a feedback control mechanism for its synthesis, which can be regulated by alterations in the adhesive properties of the substrate. The level of vinculin expression cannot simply be correlated with cellular properties such as the rate of cell growth. Vinculin synthesis was lowest in sparse cell cultures, which exhibit maximal rates of cell proliferation, whereas the synthesis was maximally stimulated in the almost arrested dense

cultures (Figs 1-3). On the other hand, vinculin synthesis was found to be lowest in the spherical non-proliferating cells in experiments where the substrate was altered (Fig. 4)^{1,2,9,18,19}.

The association of vinculin with areas of cell-substrate contact and its rapid reorganization on addition of purified growth factors to cells arrested in the G₁ phase of the cell cycle²⁰, suggests that the protein is involved in the transmission of contact-dependent, growth-related signals from the exterior. Transformation of cells by Rous sarcoma virus and other oncogenic viruses induces marked alterations in vinculin and

actin organization, concomitant with overall changes in cell shape and adhesiveness²¹⁻²³, phenomena also observed on addition of tumour promoters to cells²⁴. These observations highlight the role of cell contacts in the regulation of cell growth and suggest that proteins such as vinculin, which are involved in the construction of these cellular structures, participate in this regulation.

This work was supported in part by a grant from the Israel-USA Binational Foundation (A.B.-Z.) and by a grant from the Erna Rothstein Foundation (B.G.).

Received 31 October 1985; accepted 2 January 1986.

1. Benecke, B. J., Ben-Ze'ev, A. & Penman, S. *Cell* **14**, 931-939 (1978).
2. Ben-Ze'ev, A., Farmer, S. R. & Penman, S. *Cell* **21**, 365-372 (1980).
3. Folkman, J. & Moscona, A. *Nature* **273**, 345-349 (1978).
4. Wittelsberger, S. C., Kleene, K. & Penman, S. *Cell* **24**, 859-866 (1981).
5. Farmer, S. R., Ben-Ze'ev, A., Benecke, B. J. & Penman, S. *Cell* **15**, 627-637 (1978).
6. Ben-Ze'ev, A. in *Cell and Muscle Motility* Vol. 6 (ed. Shay, J. W.) 23-53 (Plenum, New York, 1985).
7. Bissel, M. J., Hall, H. G. & Parry, G. J. *theor. Biol.* **99**, 31-68 (1982).
8. Ben-Ze'ev, A., Farmer, S. R. & Penman, S. *Cell* **17**, 319-325 (1979).
9. Farmer, S. R., Wan, K. M., Ben-Ze'ev, A. & Penman, S. *Molec. cell. Biol.* **3**, 182-189 (1983).
10. Ben-Ze'ev, A. *J. Cell Biol.* **97**, 858-865 (1983); **99**, 1424-1433 (1984).
11. Geiger, B. *Cell* **18**, 193-205 (1979).
12. Burridge, K. & Feraisico, J. *Cell* **19**, 587-595 (1980).
13. Geiger, B., Tokuyasu, K. T., Dutton, A. H. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4127-4131 (1980).

14. Geiger, B., Dutton, A. H., Tokuyasu, K. T. & Singer, S. J. *J. Cell Biol.* **91**, 614-628 (1981).
15. Geiger, B. *Cold Spring Harb. Symp. quant. Biol.* **46**, 671-682 (1982).
16. Geiger, B., Avnur, Z., Kreis, T. E. & Schlessinger, J. in *Cell and Muscle Motility* Vol. 5 (ed. Shay, J. W.) 195-234 (Plenum, New York, 1984).
17. Ben-Ze'ev, A. in *Molecular Biology of the Cytoskeleton* (eds Borisy, G., Cleveland, D. & Murphy, D.) 435-444 (Cold Spring Harbor Laboratory, New York, 1985).
18. Johnson, L. F., Abelson, H. T., Green, H. & Penman, S. *Cell* **1**, 95-100 (1974).
19. Otsuka, H. & Moskowitz, M. J. *cell. Physiol.* **89**, 213-220 (1976).
20. Herman, B. & Pledger, W. J. *J. Cell Biol.* **100**, 1031-1040 (1985).
21. David-Pfeuty, T. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6687-6691 (1980).
22. Sefton, B. M., Hunter, T., Ball, E. H. & Singer, S. J. *Cell* **24**, 165-174 (1981).
23. Shriver, K. & Rohrschneider, L. J. *Cell Biol.* **89**, 525-535 (1981).
24. Schliwa, M., Nakamura, T., Porter, K. R. & Euteneuer, U. *J. Cell Biol.* **99**, 1045-1059 (1984).
25. O'Farrell, P. H. *J. biol. Chem.* **250**, 4007-4021 (1975).
26. Kessler, S. W. *J. Immun.* **115**, 1617-1624 (1975).

Stable transformation of maize after gene transfer by electroporation

Michael E. Fromm, Loverine P. Taylor & Virginia Walbot

Department of Biological Sciences, Stanford University, Stanford, California 94305-2493, USA

The graminaceous monocots, including the economically important cereals, seem to be refractory to infection by *Agrobacterium tumefaciens*¹, a natural gene transfer system that has been successfully exploited for transferring foreign genes into higher plants²⁻⁴. Therefore, direct transfer techniques that are potentially applicable to all plant species have been developed using a few dicot⁵⁻⁸ and monocot⁹⁻¹⁰ species as model systems. One of these techniques, electroporation, uses electrical pulses of high field strength to permeabilize cell membranes¹¹ reversibly so as to facilitate the transfer of DNA into cells^{8,12,13}. Electroporation-mediated gene transfer has resulted in stably transformed animal cells^{12,13} and transient gene expression in monocot and dicot plant cells⁸. Here we report that electroporation-mediated DNA transfer of a chimaeric gene encoding neomycin phosphotransferase results in stably transformed maize cells that are resistant to kanamycin.

The neomycin phosphotransferase II (*npt-II*) gene from the transposon Tn5 (ref. 14) is a useful dominant marker in both animal¹⁵ and plant^{3,16,17} cells when expressed using the appropriate eukaryotic regulatory signals. Expression of the *npt-II* coding region confers resistance to the antibiotics kanamycin and G418, to which plant cells are normally sensitive^{3,16,17}. We demonstrated previously that the cauliflower mosaic virus (CaMV) 35S promoter (nucleotides 7,017-7,437, see ref. 18 for nucleotide numbers) and the nopaline synthase 3' region¹⁹ were sufficient 5' and 3' regulatory signals to express the chloramphenicol acetyltransferase coding region in maize protoplasts⁸. Accordingly, the CaMV 35S promoter, the *npt-II* coding region and the nopaline synthase 3' region were used to construct pCaMVNEO (Fig. 1) to serve as a selectable marker in maize. pCaMVNEO expresses detectable *npt-II* enzyme activity 20 h after electroporation-mediated transfer into maize protoplasts (data not shown), demonstrating that it can be expressed in maize cells and therefore should confer resistance to kanamycin in transformed cells.

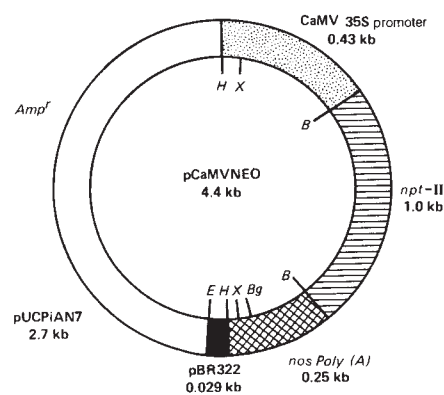


Fig. 1 Diagram of pCaMVNEO. The size in kilobase pairs is shown next to each DNA fragment composing pCaMVNEO, with the source of each fragment represented by a different background: stippled, cauliflower mosaic virus 35S promoter; hatched, neomycin phosphotransferase II; cross-hatched, nopaline synthase (*nos*) polyadenylation region; black, pBR322; white, pUCPIAN7. Relevant restriction sites are also indicated: B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; X, *Xba*I.

Methods. pUCPIAN7 contains the polylinker of PIAN7 (Biolabs) inserted into PUC8³¹. A 777-base pair (bp) *Hinc*II fragment (nucleotides 7,017-7,794, see ref. 18 for nucleotide numbers) from CaMV strain 1841 (ref. 32) containing the CaMV 35S promoter was inserted into the *Hinc*II site of pUC8. A 430-bp fragment (nucleotides 7,017-7,437) was excised with *Eco*RI and *Hph*I, then inserted into *Eco*RI, *Hinc*II cleaved pUC8. This 430-bp CaMV 35S promoter fragment was then excised as a *Bam*HI-*Pst*I fragment and inserted into pUCPIAN7, which had been double digested with *Bgl*II and *Pst*I to yield pCaMV. The 250-bp *Sau*3A fragment spanning the polyadenylation region of the *nos* gene³ was sequentially inserted into the *Bgl*II site of pUCPIAN7, then into pBR322 as a *Hind*III-*Bam*HI fragment, followed by insertion into pCaMV as an *Eco*RI-*Bam*HI fragment. This yields the plasmid pCaMVNOS that contains a unique *Bam*HI site between the CaMV 35S promoter and the *nos* polyadenylation region. A 1.0-kb *Bam*HI fragment of the *npt-II* gene spanning nucleotides 1,540-2,518¹⁴ was inserted into the *Bam*HI site of pCaMVNOS to yield pCaMVNEO. The nucleotide position corresponding to the 5' end of the CaMV 35S RNA¹⁸ (nucleotide 7,435) is located 17 bp upstream of the *Bam*HI site at the 5' end of the *npt-II* gene fragment shown.

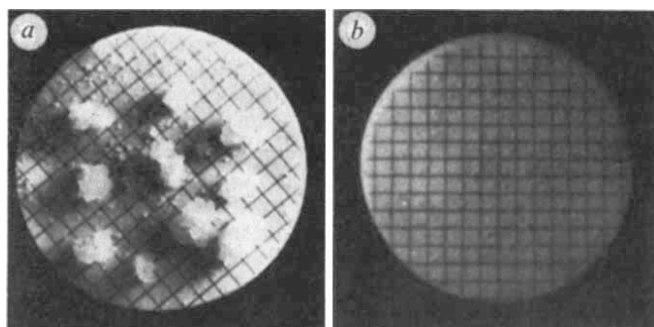


Fig. 2 Selection of kanamycin-resistant BMSI calli. The growth of BMSI microcalli (shown 8 weeks after protoplast isolation) was sensitive to kanamycin ($100 \mu\text{g ml}^{-1}$) unless the protoplasts were electroporated in the presence of pCaMVNEO DNA. The membrane filter diameter is 47 mm, with a grid spacing of 3 mm. *a*, Culture electroporated with $50 \mu\text{g ml}^{-1}$ pCaMVNEO DNA present; *b*, control culture electroporated without pCaMVNEO DNA present.

Methods. Protoplasts were isolated from 8 ml of packed cells from a rapidly growing BMSI suspension culture (obtained from P. Okubara, Stauffer Chemical Co.) by resuspending the cells in a solution containing 1% cellulase (Cellulysin, Calbiochem), 0.5% hemicellulase (Rhozyme, Genencore), 0.02% pectinase (Pectolyase Y23, Seishin Pharmaceutical Co.), 50 mM CaCl_2 , 10 mM sodium acetate, pH 5.8, 0.2 M mannitol as described elsewhere⁸. The protoplast pellet was washed once in the same solution without enzymes, then washed and resuspended in a buffer containing 2 mM phosphate, pH 7.1, 4 mM CaCl_2 , 140 mM NaCl and 0.2 M mannitol at a density of 4×10^6 protoplasts ml^{-1} . The protoplast-containing solution (0.5 ml) was mixed with 0.5 ml of the same solution containing $50 \mu\text{g}$ of supercoiled plasmid DNA, placed in a 1-ml aluminium foil-lined plastic cuvette and electroporated as described previously⁸. The electrical pulse was delivered from a 122- or 245- μF capacitor charged to 200 V. This corresponds to an electric field intensity of 500 V/cm and RC time constants of 2 and 4 ms respectively. After 10 min at 4°C and 10 min at room temperature, protoplasts were diluted with 8 ml media at room temperature containing MS salts²², 0.3 M mannitol, 2% sucrose, 2 mg l^{-1} 2, 4-D, 20% conditioned media and 0.1% low melting agarose (FMC Corporation). After 2 weeks in the dark at 26°C , medium without mannitol and containing kanamycin sulphate (Sigma) was added to give a final selective medium containing $100 \mu\text{g ml}^{-1}$ of kanamycin sulphate. Microcalli were removed from liquid selective media after 2 weeks and placed on membrane filters (GN-6, Gelman) supported on cellulose absorbant pads (Gelman) above agarose-solidified medium containing $100 \mu\text{g ml}^{-1}$ of kanamycin sulphate. Kanamycin-resistant calli appear after 1–2 weeks on the disks (5–6 weeks after protoplast electroporation). Cells electroporated without pCaMVNEO DNA present do not grow under selective conditions.

For our model DNA transformation system, we chose a Black Mexican Sweet (BMS) maize suspension culture line, BMSI (ATCC 54022), as BMS callus lines form finely divided suspension cultures²⁰ and yield protoplasts capable of reforming their cell wall and dividing²¹. BMSI maize protoplasts were isolated and electrically permeabilized in the presence of $50 \mu\text{g ml}^{-1}$ supercoiled pCaMVNEO DNA as described previously⁸ except that the electrical pulse was delivered from either 122- or 245- μF capacitors charged to 200 V. Electrical pulses of 200 V from larger capacitors gave rise to metabolically active, but non-dividing cells (data not shown).

The electroporated protoplasts were diluted in MS medium²² containing 2% sucrose, 2 mg l^{-1} 2, 4-D, 0.3 M mannitol, 20% conditioned media and 0.1% low melting agarose. Approximately 1% of the electroporated protoplasts reformed their cell wall and divided after 1 week, compared with 5% for non-electroporated protoplasts. Two weeks after electroporation, fresh medium containing kanamycin sulphate, was added to the

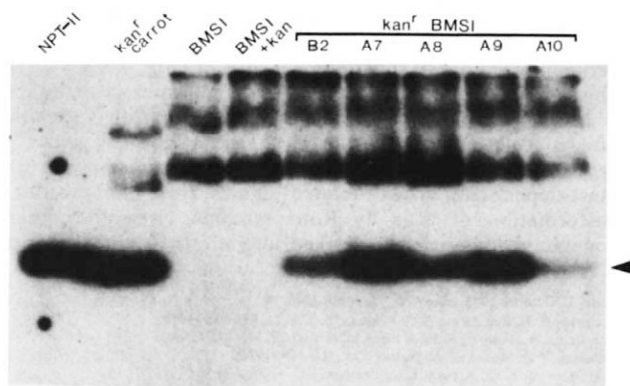


Fig. 3 Neomycin phosphotransferase II assay of extracts from kanamycin-resistant BMSI calli. Extracts ($50 \mu\text{l}$) prepared from kanamycin-resistant BMSI calli and untreated control BMSI calli were electrophoresed through a 10% polyacrylamide gel and assayed for *npt-II* activity *in situ* using kanamycin and [γ - ^{32}P]ATP as substrates²³. Extracts from an *E. coli* strain containing the *npt-II* gene and from an *A. tumefaciens*-transformed carrot callus containing a chimaeric *npt-II* gene³ serve as positive controls for *npt-II* activity and gel position. The kanamycin-resistant BMSI calli contain *npt-II* activity that is not present in untreated control BMSI callus (BMSI lane). BMSI callus maintained on medium containing $100 \mu\text{g ml}^{-1}$ of kanamycin sulphate for 5 days before analysis does not produce *npt-II* activity (BMSI + kan lane). Callus line BMSI-B2 was obtained in a separate experiment from that giving rise to callus lines BMSI-A7, -A8, -A9 and -A10. The arrow indicates the position of the *npt-II* activity.

Methods. Calli (50–100 mg) were disrupted at 0°C with a ground glass mortar and pestle in an equal volume of buffer containing 62 mM Tris, pH 6.8, 10% glycerol, 5% β -mercaptoethanol, 250 $\mu\text{g ml}^{-1}$ phenylmethylsulphonyl fluoride, 130 $\mu\text{g ml}^{-1}$ leupeptin and 50 $\mu\text{g ml}^{-1}$ bromophenol blue. Aliquots ($50 \mu\text{l}$) of each extract were electrophoresed through a polyacrylamide gel and assayed for *npt-II* activity as reported elsewhere²³.

dividing cells to give a final kanamycin sulphate concentration of $100 \mu\text{g ml}^{-1}$. After 2 weeks in this medium, the microcalli were placed on a membrane filter disk lying on a cellulose absorbant pad above agarose-solidified media containing $100 \mu\text{g ml}^{-1}$ kanamycin sulphate. During the first 2 weeks on the disks, kanamycin-resistant colonies grew from cells electroporated with pCaMVNEO DNA, while no colonies grew from cells electroporated without pCaMVNEO DNA (Fig. 2). Colonies of 0.5 cm diameter were then placed directly on agarose-solidified medium containing $100 \mu\text{g ml}^{-1}$ kanamycin sulphate and reached a weight of 3–6 g by 12 weeks after the electroporation treatment. Eight independent preparations of BMSI protoplasts have yielded kanamycin-resistant colonies when electroporated in the presence of pCaMVNEO DNA. The highest frequency observed was 161 kanamycin-resistant colonies per 2×10^6 initial protoplasts. No kanamycin-resistant colonies have appeared in control cells.

Kanamycin-resistant BMSI calli from two independent experiments were assayed for the presence of *npt-II* enzyme by electrophoresing cell extracts through a 10% polyacrylamide gel and detecting enzyme activity by an *in situ* assay using kanamycin and [γ - ^{32}P]ATP (ref. 23). The kanamycin-resistant calli contain a phosphotransferase activity not present in untreated kanamycin-sensitive calli (Fig. 3). This new phosphotransferase activity co-migrates with the *npt-II* protein present in extracts from bacteria containing the *npt-II* gene (Fig. 3, *npt-II* lane) and from carrot callus containing a modified T-DNA region of *A. tumefaciens* that carries a chimaeric *npt-II* gene³ (Fig. 3, *kan^r carrot* lane). Untreated BMSI callus does not contain this activity even when placed on kanamycin-containing

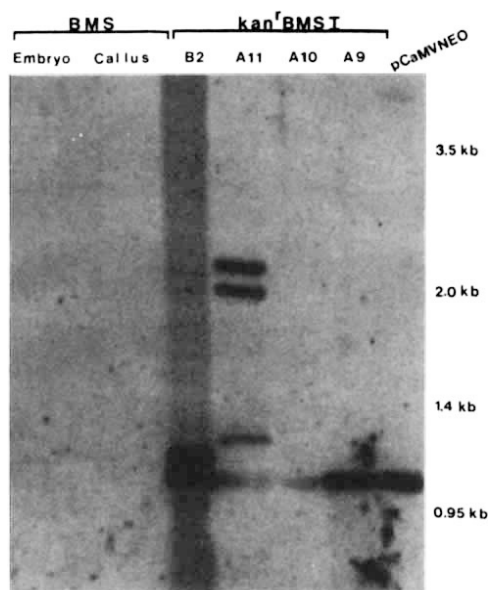


Fig. 4 Southern blot analysis of DNA from BMSI calli. DNA was isolated from untreated control BMSI callus, BMS embryos and kanamycin-resistant BMSI calli from two independent electroporation experiments (BMSI-A and BMSI-B). The DNA ($\sim 5 \mu\text{g}$ per lane) was electrophoresed through a 1% agarose gel, transferred to a nylon membrane (Genetran Plasco) and the blots were hybridized³³ with the ³²P-labelled²⁶ 1.0-kb *Bam*HI fragment from pCaMVNEO (Fig. 1) which contains the *npt-II* gene. Callus and embryo DNA was isolated by a protocol used for single maize seedlings³³. *Bam*HI-digested DNA from kanamycin-resistant calli (lanes BMSI-A9, -A10, -A11 and -B2) contains the diagnostic 1.0-kb *npt-II* fragment that co-migrates with the pCaMVNEO *Bam*HI *npt-II* fragment (pCaMVNEO lane, present at 5 copies per maize haploid genome). *Bam*HI-digested DNA from BMS embryos (BMS embryo lane) or two control BMSI calli (BMS callus lanes) does not hybridize to the *npt-II* probe. The autoradiograph exposure was 11 days at -80°C using Kodak XAR-5 film with Lightning Plus intensifying screen (Dupont).

media for 5 days before the assay (Fig. 3, BMSI+kan lane). The more slowly migrating bands present in both the carrot and BMSI callus extracts have been observed for several plant species and are thought to represent protein kinases²⁴.

Southern blot analysis²⁵ of DNA from the kanamycin-resistant BMSI calli and from the kanamycin-sensitive BMSI callus control was performed to confirm that the new phosphotransferase activity was correlated with the presence of the introduced CaMVNEO gene. The 1.0-kilobase (kb) *Bam*HI *npt-II* gene fragment (Fig. 1) was isolated and labelled radioactively²⁶ for use as a hybridization probe to detect pCaMVNEO sequences (Fig. 4). DNA from untransformed BMSI callus or BMS embryos does not contain DNA sequences that hybridize to the *npt-II* probe (Fig. 4, BMS callus and BMS embryo lanes). *Bam*HI-digested DNA from kanamycin-resistant BMSI transformants shows hybridization to the expected 1.0-kb *npt-II* fragment (Fig. 4, kan^R BMSI lanes) that co-migrates with the 1.0-kb *npt-II* fragment from *Bam*HI-digested pCaMVNEO control DNA (Fig. 4, pCaMVNEO lane, present at an equivalent of 5 copies per maize haploid genome). Thus, the kanamycin-resistant BMSI calli each contain a 1.0-kb *npt-II* fragment from pCaMVNEO. Transformants BMSI-A9 and BMSI-A10 apparently contain only intact copies, while transformants BMSI-B2 and BMSI-A11 contain additional, rearranged copies of the CaMVNEO gene (Fig. 4). The four transformants examined contain one to five copies of the CaMVNEO construct per haploid genome, after correcting for variation in the amount

of DNA present in each lane by hybridizing the blot with the maize *Bronze* gene²⁷ as probe.

Undigested DNA from the kanamycin-resistant calli probed with the 1.0-kb *npt-II* radioactive fragment shows hybridization only to the DNA of high relative molecular mass ($> 20 \text{ kb}$, data not shown). There is no hybridizing material in the area where the circular pCaMVNEO plasmid migrates. These results indicate that the pCaMVNEO sequences are probably integrated into maize chromosomes. We verified that BMSI control and transformed cells are derived from maize BMS plants by alcohol dehydrogenase (ADH) isozyme and Southern blot analyses. BMSI callus cells contain ADH genes that produce ADH proteins with electrophoretic mobilities²⁸ identical to the *Adh1-F*- and *Adh2*-encoded ADH proteins from BMS plants (data not shown). We further verified by Southern blot analysis²⁵ that DNA from control and transformed BMSI cells and from BMS plants contain *Adh1* genes with identical *Bam*HI, *Bgl*II and *Hind*III restriction site locations (data not shown).

These experiments demonstrate that the electroporation method of gene transfer into plant protoplasts⁸ yields transformed maize cells containing and expressing the CaMVNEO gene. The expression of the *npt-II* coding region confers resistance to kanamycin and is useful as a dominant selectable marker in maize cells, as has been found in other plant species^{3,16,17}. Up to 1% of the dividing protoplasts become kanamycin resistant when electroporated in the presence of pCaMVNEO DNA, a frequency which compares favourably with those of other direct DNA transformation systems (10^{-6} to 10^{-2} of dividing cells)^{5,6,9,10,34}. Direct DNA transformation techniques are advantageous for cereals because of the apparent host-range limitations of *A. tumefaciens* to dicots¹ and a few non-graminaceous monocots^{29,30}.

We thank Judy Callis and Carol Rhodes for helpful discussions, Thanh Huynh for the DNA isolation protocol and Patricia Okubara (Stauffer Chemical Co.) for the gift of the BMSI cell line. The technical assistance of Anthony Oro and Chris Warren is gratefully acknowledged. This work was supported by NIH grant GM32422 (V.W.), a Pioneer Hi-bred International fellowship (M.F.) and a predoctoral traineeship from USPHS CA09302 (L.P.T.).

Received 28 October; accepted 17 December 1985.

- De Cleene, M. & De Ley, J. *Bot. Rev.* **42**, 389-466 (1976).
- Caplan, A. *et al. Science* **222**, 815-821 (1983).
- Frabley, R. T. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 4803-4807 (1983).
- Zambryski, P., Herrera-Estrella, L., De Block, M., Van Montagu, M. & Schell, J. in *Genetic Engineering Vol 6* (eds Hollaender, A. & Setlow, J.) 253-278 (Plenum, New York, 1984).
- Krens, F. H., Molendijk, L., Wullems, G. J. & Schilperoort, R. A. *Nature* **296**, 72-74 (1982).
- Hain, R. *et al. Molec. gen. Genet.* **199**, 161-168 (1985).
- Paszkowski, J. *et al. EMBO J.* **3**, 2717-2722 (1984).
- Fromm, M., Taylor, L. P. & Walbot, V. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5824-5828 (1985).
- Lörz, H., Baker, B. & Schell, J. *Molec. gen. Genet.* **199**, 178-182 (1985).
- Potrykus, I. *et al. Molec. gen. Genet.* **199**, 183-188 (1985).
- Zimmerman, U. & Vienken, J. *J. Membrane Biol.* **67**, 165-182 (1982).
- Potter, J., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161-7165 (1984).
- Neumann, E., Schaefer-Ridder, M., Wang, Y. & Hofschneider, P. H. *EMBO J.* **1**, 841-845 (1982).
- Beck, E., Ludwig, G., Auerswald, E. A., Reiss, B. & Schaller, H. *Gene* **19**, 327-336 (1982).
- Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327-341 (1982).
- Herrera-Estrella, L. *et al. EMBO J.* **2**, 987-995 (1983).
- Bevan, M. W., Flavell, R. B. & Chilton, M.-D. *Nature* **304**, 184-187 (1983).
- Hohn, T., Richards, K. & Lebeurier, G. *Curr. Topics Microbiol. Immun.* **96**, 193-236 (1982).
- Bevan, M., Barnes, W. M. & Chilton, M.-D. *Nucleic Acids Res.* **11**, 369-385 (1983).
- Sheridan, W. F. *J. Cell Biol.* **67**, 396a (1975).
- Chourey, P. S. & Zurawski, D. B. *Theor. appl. Genet.* **59**, 341-344 (1981).
- Murashige, T. & Skoog, F. *Physiol. Pl.* **15**, 473-497 (1962).
- Reiss, B., Sprengel, R., Will, H. & Schaller, H. *Gene* **30**, 211-218 (1984).
- Schreier, P. H., Seftor, E. A., Schell, J. & Bohnert, H. J. *EMBO J.* **4**, 25-32 (1985).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
- Federoff, N. V., Furtek, D. B. & Nelson, O. E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3825-3829 (1984).
- Freeling, M., Cheng, D., S.-K. & Alleman, M. *Dev. Genet.* **3**, 179-196 (1982).
- Hooykass-Van Slooten, G. M. S., Hooykass, P. J. J. & Schilperoort, R. A. *Nature* **311**, 763-764 (1984).
- Hernalsteens, J.-P., Thia-Toony, L., Schell, J. & Montagu, M. *EMBO J.* **3**, 3039-3041 (1984).
- Vieira, J. & Messing, J. *Gene* **19**, 259-268 (1982).
- Gardner, R. C. *et al. Nucleic Acids Res.* **9**, 2871-2887 (1981).
- Rivin, C. J., Cullis, C. A. & Walbot, V. *Genetics* (in the press).
- Shillito, R. D., Saul, M. W., Paszkowski, J., Müller, M. & Potrykus, I. *Bio/Technology* **3**, 1099-1103 (1985).

Lack of antibodies to adult T-cell leukaemia virus and to AIDS virus in Israeli Falashas

ANTIBODIES to the virus involved in the development of adult T-cell leukaemia (adult T-cell leukaemia virus, ATL/HTLV-I) were recently reported in 88 of 233 Falasha immigrants to Israel¹. Ten coded sera from these individuals, including 7 sera previously found to be positive, were made available to us to be tested for antibodies to ATL/HTLV-I. Using immunofluorescence and immunoperoxidase tests, we found that none of these sera contained antibodies to the virus. The same 10 coded sera were also sent to Dr Hinuma's laboratory in Kyoto, Japan, where they were tested by immunofluorescence, but again none was found to have antibodies to ATL/HTLV-I (Table 1). A further 80 coded sera from Israel were also tested in Cambridge, UK, for antibodies to ATL/HTLV-I (Table 1). None of these 80 sera, including 55 from the Falashas

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

human retrovirus, HTLV-1, from cultured cells of human acute myeloid leukaemia² which subsequently proved to be caused by laboratory contamination by two monkey viruses^{3,4}, there were claims that

distinct from oncoviridae (ATLV/HTLV-I)¹⁰. Because LAV is thought to be of African origin, we have also tested the 55 sera from the Falashas for the AIDS virus, but none contained antibodies to LAV.

Perhaps the rapid pace of research in this area encourages the emergence of spurious claims and the publication of premature data, lacking confirmation and requiring subsequent rebuttal.

A. KARPAS*
S. MAAYAN†
R. RAZ†

* *Department of Haematological Medicine, University Clinical School, Hills Road, Cambridge CB2 2QL, UK*
† *Hadassah University Hospital, Jerusalem and The Emek Hospital, Afula, Israel*

Received 5 August 1985.

Table 1 Tests for anti-ATLV, HTLV-I and LAV antibodies

Sera	Cambridge		Kyoto	Haifa
	ATLV/HTLV-I	LAV	ATLV/HTLV-I	ATLV/HTLV-I
Japanese ATL patients	5/5	0/5	5/5	ND
British West Indians with ATL	2/2	0/2	2/2	ND
British West Indians without ATL	0/71	ND	ND	ND
Falashas	0/10	0/10	0/10	7/10
	0/55	0/55	ND	Not known
British homosexuals with and without AIDS	0/228	63/228	ND	ND
US AIDS patients	0/35	35/35	0/35	ND

Values shown are number of sera positive for the respective antibodies/number tested. Immunofluorescence tests were performed in the laboratory of Professor Y. Hinuma, as described previously¹¹. In Cambridge, both immunofluorescence and immunoperoxidase methods were used as described previously.¹² The Karpas 45 T-cell line infected with the Cambridge isolate of LAV was used to screen for antibodies to the AIDS virus. As shown above, only the Japanese and British patients with ATL had antibodies to ATL/HTLV-I. None of the 1017 British and US individuals with antibodies to the AIDS virus had detectable antibodies to ATL/HTLV-I. ND, not done.

that arrived recently in Israel, contained antibodies. No documented case of adult T-cell leukaemia (ATL) in immigrant Falashas has yet been reported.

A difference in sensitivity between various tests cannot explain so marked a discrepancy between the concordant negative findings in Kyoto and Cambridge on the one hand and the 37% positivity recorded by Ben-Ishai *et al.*¹ on the other. Unfortunately, misleading reports on viral serology have been common in recent years. Similar discrepancies have, for example, been reported in relation to serological studies of retroviruses in human leukaemia and also acquired immune deficiency syndrome (AIDS). Following reports of the isolation of a putative

a highly sensitive serological test could detect antibodies to one of the involved monkey viruses in 100% (39 of 39) of human sera and antibodies to the other virus in 49% of human sera⁵. These claims were later refuted and the reaction shown not to be serologically specific^{6,7}.

Again, publications as late as 1983 repeatedly claimed the widespread presence of antibodies to ATL/HTLV-I in AIDS patients as well as in haemophiliacs⁸, claims which could not be confirmed⁹. We now know that the retrovirus that causes AIDS, lymphadenopathy-associated virus (LAV), belongs to a different group of viruses, the lentiviridae, which are morphologically, serologically, biologically and genetically

Lack of HTLV-I antibodies in Africans

HUMAN T-cell leukaemia virus type I (HTLV-I), the aetiological agent of adult T-cell leukaemia-lymphoma (ATLL), is endemic in south-west Japan and the Caribbean basin¹. Within the Japanese endemic regions, the prevalence of serum antibodies to HTLV-I varies in different localities from 6 to 37% of healthy adults tested². HTLV-I is thought to be widespread in Africa, as antibodies cross-reacting with HTLV-I antigens have been reported in sera taken from Africans in many different countries throughout the continent^{3,4}. A recent report from Ben-Ishai *et al.*⁵ recorded a particularly high prevalence of anti-HTLV-I (36.9%) in Ethiopian Jewish immigrants (Falashas) in Israel.

We have also screened sera from Africans for anti-HTLV-I, and in contrast to the reports cited above, we failed to detect reactive sera (Table 1). Our screening test was a competition assay⁶, using HTLV-I-infected C91/PL cell lysate as the source of antigen. Antibodies used for the competition assay were reference human

anti-HTLV-I sera from British subjects conjugated with either ^{125}I or horseradish peroxidase. The test gave strongly positive results with sera from ATLL patients and from 4% of UK subjects of West Indian origin (Table 1).

The competitive anti-HTLV-I test is probably more specific than the immunoglobulin-binding enzyme-linked immunosorbent assays (ELISAs) previously used³⁻⁵, as cross-reaction with

I, as well as for anti-HTLV-III (refs 10-12), particularly among malarial patients with circulating immune complexes^{10,13}. Clearly, there is a need to investigate these discrepancies, including exchange of sera between laboratories using different tests, and this is underway.

We thank J.-M. Bechet, D. Catovsky, R. Downing, H. Hoshino, C. Peckham and D. Serwadda for supplying sera, and P. Clapham and J. Weber for conducting neutralization and immunofluorescence tests.

Table 1 Absence of antibodies specific to HTLV-I in sera from African populations

Source	No. positive/ no. tested for HTLV-I antibodies
Ethiopian immigrants to Israel	0/259
Uganda	0/400
Malawi	0/366
Ivory Coast	0/200
UK West Indians	19/480
UK ATLL patients	12/12
Japanese ATLL patients	10/10

HTLV-II is minimal⁶. In view of the negative results with the competitive ELISA, we also screened the 259 Ethiopian Jewish sera by the Biotest ELISA alkaline phosphatase anti-HTLV-I test. The results yielded a single gaussian distribution of light absorbance (A) for the alkaline phosphatase reaction from 0.5 to 2.5 U. Only two sera had $A > 2.0$ and 12 had $A > 1.5$, which was taken as the cut-off for positive scores. The two sera with the highest reaction were tested further for anti-HTLV-I by indirect immunofluorescence of C91/PL cells⁷, by syncytium inhibition⁷ and by pseudotype neutralization⁸ tests, but were negative by all three methods.

The Ethiopian Jewish sera were derived largely from adults hospitalized for suspected or proven tuberculosis. The patients came from the Gonder region of Ethiopia, as did those studied by Ben-Ishai *et al.*⁵ The sera of most of these patients were also positive for hepatitis B virus markers. Our sample therefore comprises patients harbouring other infectious agents. In the Japanese endemic region, patients hospitalized in infectious-disease wards have a higher prevalence of anti-HTLV-I than do healthy controls or surgical patients⁹, so the absence of anti-HTLV-I in our sample of Ethiopian Jews remains puzzling.

Our findings differ from previous surveys of anti-HTLV-I in Africans³⁻⁵. It could be argued that our sera were all taken by chance from uninfected subjects within endemic regions, but gross sampling errors seem unlikely with some 1,200 sera tested from four African countries (Table 1). This raises the question of whether African sera tend to give spuriously positive results with certain anti-globulin ELISA tests for anti-HTLV-

R. A. WEISS*
R. CHEINGSONG-POPOV*
S. CLAYDEN*
S. PEGRAM*
R. S. TEDDER†
A. BARZILAI‡
E. RUBENSTEIN‡

* Institute of Cancer Research,
Fulham Road,
London SW3 6JB, UK

† Section of Virology,
Middlesex Hospital
Medical School,
London W1N 8AA, UK

‡ Unit of Infectious Diseases,
Chaim Sheba Medical Centre,
Sackler School of Medicine,
Tel-Hashomer, Israel

Received 14 November 1985.

- Vogt, P. K. (ed.) *Curr. Topics Microbiol. Immun.* **115**, 1-266 (1985).
- Hinuma, Y. *et al. Int. J. Cancer* **29**, 631-635 (1982).
- Saxinger, W. *et al. Science* **225**, 1473-1476 (1984).
- Hunsmann, G. *et al. Med. Microbiol. Immun.* **173**, 167-170 (1984).
- Ben-Ishai, Z. *et al. Nature* **315**, 665-666 (1985).
- Tedder, R. S. *et al. Lancet* **ii**, 125-128 (1984).
- Nagy, K., Clapham, P., Cheingsong-Popov, R. & Weiss, R. A. *Int. J. Cancer* **32**, 321-328 (1983).
- Clapham, P., Nagy, K. & Weiss, R. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3083-3086 (1984).
- Essex, M., McLane, M. F., Tachibana, N., Francis, D. P. & Lee, T. H. in *Human T-Cell Leukemia/Lymphoma Virus* (eds Gallo, R. C., Essex, M. & Gross, L.) 355-362 (Cold Spring Harbor Laboratory, New York, 1984).
- Biggar, R. J. *et al. Lancet* **ii**, 520-524 (1985).
- Hunsmann, G., Schneider, J., Wendler, I. & Fleming, A. F. *Lancet* **ii**, 952-953 (1985).
- Serwadda, D. *et al. Lancet* **ii**, 849-852 (1985).
- Mohammed, I. J. *Infect.* **4**, 97-104 (1982).

BEN ISHAI *ET AL.* REPLY—Weiss *et al.* (above) did not find antibodies specific to HTLV-I in sera from 1,225 serum samples from donors of different African countries. As they state, others have found antibodies to HTLV-I in sera of Africans from different locations, using enzyme immunoassays and immunofluorescence.

We have recently received 127 serum samples donated by Ethiopian immigrants to Israel. These same samples had been tested by Weiss *et al.* and found to be negative for antibodies to HTLV-I. In our ELISA test for HTLV-I antibody all the sera were negative. Two of these sera that gave higher absorbance readings were re-tested in a confirmatory assay by Western blots and found to be negative.

Among the original Falasha serum samples obtained in 1983 (ref. 1), 37% showed reactivity in the ELISA test for antibody to HTLV-I. Fourteen of the positive samples that are still available from the original (1983) shipment have been re-tested by Western blotting. All 14 sera precipitated a protein of relative molecular mass 53,000, presumably a precursor to the p24 gag protein of HTLV-I.

HTLV-I has been isolated from individuals in the United States, Japan, the Caribbean basin, Africa, Latin America and Europe. Three isolates of HTLV-I have been obtained from Israelis². Saxinger *et al.*³ have suggested that human retroviruses cross-reactive to HTLV-I are present in various populations. Antibodies found in Falasha sera¹ may reflect such cross-reactivity. We fully agree with Weiss *et al.* that there is a need to determine the nature of these immunological cross-reactivities, and to isolate the virus from antibody-positive individuals.

ZVI BEN ISHAI*
MARTIN HAAS†
FRED C. JENSEN‡

* Rambam Medical Centre,
Haifa 32096, Israel

† Cancer Center, University of California,
San Diego, La Jolla,
California 92093, USA

‡ Cytotech Inc, San Diego,
California 92121, USA

Received 10 December 1985.

- Ben Ishai, Z. *et al. Nature* **315**, 665-666 (1985).
- Robert-Guroff, M. *et al.* in: *Current Topics in Microbiology and Immunology: Human T-Cell Leukemia Virus* (ed. Vogt, P. K.) 7-31 (Springer, Berlin, 1985).
- Saxinger, W. *et al. Science* **225**, 1473-1476 (1984).

Dating volcanic ash by ESR

RECENTLY, Imai *et al.*¹ have described a novel application of ESR (electron spin resonance) techniques to the dating of Quaternary volcanic products. Specifically, they have selected the aluminium-centre ESR signal (at 77 K) in quartz, plagioclase and glass, to increase the signal-to-noise ratio compared with the signal from the Ge and E' centres, and have reported several ESR apparent ages. Because of the encouraging agreement of one result obtained with the glass component, they have applied this technique to the glass component of six tephra of unknown age, producing results (0.5-0.8 Myr) in accord with the stratigraphic sequence.

Notwithstanding the potential usefulness of this new application of the Al-centre signal for dating and for 'fingerprinting' of tephra layers, these dating results should be accepted only with caution because several important potential sources of error have not been discussed

or properly handled. These factors are, in order of significance: (1) linear extrapolation of ESR signals in ESR versus dose plots, (2) anomalous fading, (3) sensitivity of the ESR signals to light, (4) anisotropy of the glass grains, (5) past water content of the tephra layers, and (6) the α efficiency factor for glass.

First, for the three samples of known age (no statement was made for the others), linear extrapolation of ESR growth curves was used to estimate the "total dose" (more properly, the equivalent dose). Although this is reasonable for the plagioclase, it is not for the quartz and glass where the extrapolation is 2–3 times the applied dose. The need to consider nonlinearity (due to saturation of traps) in such growth curves has been amply considered elsewhere^{2,3}. The limited range of applied doses chosen for the quartz and glass of Imai *et al.* cannot yield acceptable extrapolations, whether linear or nonlinear curve fitting is used.

Related to this problem are the estimates of mean lifetimes of traps, which were used by Imai *et al.* to imply acceptability of their calculated ages. There are two inconsistencies. First, their thermal annealing curve for quartz suggests a trap depth of ~1.5 eV, rather than 2.4 eV, as stated in ref. 1. With their stated trap depths and frequency factor, the calculated mean lifetimes (at 10 °C) are >10²¹ Myr, not 2–4 Myr, 4 Myr and >10³ Myr (for quartz, plagioclase and glass) as claimed by Imai *et al.* Furthermore, their thermal annealing data for volcanic glass suggests a distribution of trap depths, as expected for glass⁴, rather than a single trap depth as assumed by Imai *et al.* Second, if the quartz mean lifetimes are 2–4 Myr, calculated ages greater than about 1 Myr would underestimate the true age by >20% because of continuous trap emptying at ambient temperatures. This effect diminishes if the calculated quartz age of 1.1 Myr is decreased, for example by avoiding a linear extrapolation.

The lack of any comment by these authors on the effects of anomalous fading, exposure of samples to light, anisotropy of the glass grains, and past pore-water content in the tephra deposits is unsettling. The thermoluminescence (TL) signal from clear volcanic glass does not exhibit anomalous fading⁵, but it is well known that volcanic airfall glass shards larger than ~100 μ m commonly contain crystallites and inclusions of minerals such as feldspars. Volcanic feldspars are notorious for the instability of their TL signals. No tests for such instability in the ESR signals from these materials

have been reported. Such fading can lead to severe underestimation of ages, and in combination with linear extrapolation can produce fortuitously 'correct' ESR ages. Furthermore, the TL signals of quartz, plagioclase and glass are easily reduced by visible light during collection and handling. It is reasonable to expect similar effects for ESR signals but these have not been discussed by Imai *et al.* Finally, the mean past pore-water content for terrestrial tephra generally can be as great as 50% ($\pm 20\%$); this implies a 20% error in the ESR apparent age. Because Imai *et al.* have not mentioned water content, they must have assumed it to have been <10%, which is clearly unrealistic.

A final (and less significant) point to be made is that the α efficiency factor of 0.04 (k -value) chosen by Imai *et al.* for volcanic glass is inaccurate. This choice comes from ref. 2, but in ref. 2 this value is only a guess for a tektite glass, assuming it equals that for a quartz (!) sample. Elsewhere⁵ I have measured α efficiency factors for several glass-rich and one glass-pure sample, yielding α -values (= k -value at 3.7 MeV) of 0.08–0.13, average 0.1. Without direct measurement on a new material, it is always preferable to choose a value of 0.1, assuming uniformity of α dose.

In conclusion, because of neglect of most of these sources of error by Imai *et al.*, caution should be exercised in assessing their age results.

G. W. BERGER

Department of Physics,
Simon Fraser University,
Burnaby, British Columbia,
Canada V5A 1S6

1. Imai, N., Shimokawa, K. & Hirota, M. *Nature* **314**, 81–83 (1985).
2. Hennig, G. J. & Grün, R. *Q. Sci. Rev.* **2**, 157–238 (1983).
3. Wintle, A. G. & Huntley, D. J. *Q. Sci. Rev.* **1**, 31–33 (1982).
4. Berger, G. W. & Huntley, D. J. *Dating volcanic ash by thermoluminescence*, PACT J., Council of Europe **9**, 581–592 (1983).
5. Berger, G. W. *J. Volcanol. geotherm. Res.* **25**, 333–347 (1985).

IMAI *ET AL.* REPLY—Following on the comments of Berger, we have recalculated several values and, on the basis of this we do not believe that the error mentioned by Berger is sufficient to prevent the application of our technique to volcanic ash¹. Nonlinear fitting to the saturation curve yields 0.95 Myr for quartz (TM1) and 0.31 Myr for volcanic glass (TE5). Linear fitting yields an error of 15% for quartz and 5% for volcanic glass.

In our paper¹ we assigned E too high a value; as pointed out by Berger, the correct value is 1.4–1.5 eV. The mean lifetime was obtained assuming a simple process for the decay of the trap, and no other complicated structure was considered.

There is no anomalous fading. The ESR signals of quartz, plagioclase and volcanic glass are quite stable and the intensity remains constant for over 1 yr. The signal is also highly stable when exposed to light. The intensity change of the signal was within the experimental error when 0.1 g of sample was exposed to a 40-W tungsten lamp placed 10 cm from the sample for 27 h. Illumination with ordinary fluorescent light during sample treatment and ESR measurement also has no effect. We confirmed that the separated volcanic glass was pure and homogeneous by microscopic examination.

The effect of water was considered. There was a high water content (26%) only in the case of the TE5 ash, which results in an error of 22% for the ESR age². However, the errors for the ages of the other ashes lie within 3%, because the water contents of volcanic ashes from the Takio Formation were 1–2.7% and those derived from rocks (TM1) and (TTS34) were <1%. The water content should change, on the whole from the time of deposition to the present, thus it is not acceptable to regard the present water content as being representative of the water content through geological time.

The value of 0.1 for the k factor is, we think, an improvement. When this value is used, the ages of volcanic ashes 1–6 change from 0.49–0.82 Myr to 0.42–0.72 Myr, which represents a change of 12–15%.

Calculation of ESR ages is based on several assumptions, which sometimes lead to error. However, this does not reduce the validity of our technique. At present, comparison of the ESR age with other dating methods is the most practical means of assessing the accuracy of the age, despite the uncertainty caused by potential errors.

NOBORU IMAI

KOICHI SHIMOKAWA

Geological Survey of Japan,
1-1-3 Higashi, Yatabe, Tsukuba,
Ibaraki 305, Japan

MICHIO HIROTA

Meteorological Research Institute,
1-1 Nagamine, Yatabae, Tsukuba,
Ibaraki, Japan

1. Imai, N., Shimokawa, K. & Hirota, M. *Nature* **314**, 81–83 (1983).
2. Hennig, G. J. & Grün, R. *Q. Sci. Rev.* **2**, 157–238 (1983).

Intelligent investigations

P.E. Bryant

Theories of Development: Concepts and Applications, 2nd Edn. By William C. Crain. Prentice-Hall: 1985. Pp.306. Pbk \$20.95, £22.40.

Human Intelligence: Perspectives and Prospects.

By Robert Kail and James W. Pellegrino.

W.H. Freeman: 1985. Pp.224. Hbk \$19.95, £22.95; pbk \$12.95, £14.95.

Human Abilities: An Information-Processing Approach. Edited by R.J. Sternberg.

W.H. Freeman: 1985. Pp.259. Hbk \$19.95, £23.95; pbk \$13.95, £15.95.

PSYCHOLOGISTS began to construct tests of intelligence about 90 years ago, and from the start they had at least one certainty to help them choose the right material. They were sure that intelligence (whatever it is) grows during childhood, and therefore that any decent test would have to be more difficult for younger children to do than for older ones. So these pioneers set about finding tasks which tell younger and older children apart. An incidental consequence was that, as well as finding out how to measure intelligence, they launched the subject of developmental psychology. Both the initial uncertainty about the exact nature of intelligence and the certainty that it does at least develop in childhood are with us to this day, and they play a large role in each of these three books.

William Crain's is about the developmental side. It is the second edition of an extremely useful book which gives a clear account of many of the most important theories of intellectual and social development. The main distinction made is between theories which stress the possibility that internal forces control a child's development and those which argue for the overriding importance of the effects of experiences in the environment.

Crain spreads his net widely and some of his most interesting material is about theorists who usually do not figure much in books about developmental psychology. Thus he writes about Locke, Rousseau and also Montessori, and shows the important role their ideas had in the history of child psychology. The book has its limitations, however. It is about theories and has very little room for empirical evidence, which means that at times it is rather too uncritical. There is also one odd omission. Crain ignores the work of Vygotsky and of other Russian developmental psychologists, a particularly surprising gap because inclusion of the Russian contributions would have given the book a far better balance. As it is, most of the theories described are about the importance of internal forces in development; the Russians have more to say about environmental effects. Nevertheless *Theories of Development* is clear, enthusiastic and erudite, and it should

help beginners and experts alike.

Though many of the ideas described by Crain are about the growth of intelligence in childhood, they do not tell us much about the psychology of adult intelligence. Theories about this vexed question have undergone a marked change during the past decade, as the book written by Robert Kail and James Pellegrino and the one edited by Robert Sternberg show. Until ten years ago, such theories tended to come in one or other of two distinct forms, either dealing with differences between people (based on the results of psychometric tests) or with the nature of intelligence in general. These two kinds of theory have had very different histories.

Those based on individual differences did not prosper. The tests were, and still are, a resounding practical success; but the theories which were derived from the results were all rooted in factor analysis, a statistical technique which allows one to make claims about the ingredients of intelligence ("intelligence consists of abilities a, b and c") but not about the way the system works. These static theories failed to convince many people and eventually went into a sharp well deserved decline.

It was virtually the opposite story with the theories about the mechanisms underlying apparently intelligent acts like remembering, learning, reasoning and problem solving. Such theories, which are commonly called information-processing theories, did not have much practical impact. But they did two things — one was to provide us with the viable account of the workings of intelligent behaviour, and the other was to produce a set of remarkably successful experimental paradigms.

The big change, which is charted by both Kail and Pellegrino and the contributors to Sternberg's book, arrived when these two separate traditions merged and people began to apply information-processing theories and the experimental paradigms that went with them to individual differences. The two books take much the same point of view and deal with much the same material when they describe this development. Both conclude that there is very good reason for using models of information processing to account for human intelligence.

How well do they make their case? The main way that both books set about doing so is to cite evidence for a strong relationship between a subject's performance in the experimental paradigms of information-processing and that in particular types of intelligence tests. To take two examples, skill at judging whether two words represent things in the same taxonomic category (a common information-processing task) is related to verbal intelligence scores, while success at working out whether two subjects presented in different rotational planes are identical objects or not (another information-processing task) is related to performance in spatial intelligence tests. Note the specificity of these connections.

There are, however, some gaps in the evidence. Many of the relationships reported in these two books are not nearly specific enough. For example Earl Hunt, one of the contributors to Sternberg's book, claims a strong relationship between performance in various verbal information-processing experiments and in verbal intelligence tests; but he fails to point out — and it is a very serious omission — that in many of the cases that he describes the relationship is much too broad. Nonetheless, the two books do manage to make a convincing case for using information-processing ideas in theories about human intelligence, and both of them do so in a clear and simple way which students will find extremely helpful.

Despite this brave, new impetus we cannot yet say exactly what human intelligence is. In fact the single certainty that we do have is the one with which the subject began — that intelligence increases during childhood. All three books make much of this plain fact. Kail and Pellegrino, and Crain, give good, clear accounts of the main theories of the development of intelligence. The developmental theme also looms large in Sternberg's book, most notably in a first-rate chapter by Campione, Brown and N. Bryant. Among other things, these authors describe some evidence of marked differences between younger and older children whose absolute scores (and therefore mental ages) are the same; the younger ones form rules and generalize from one problem to another much more efficiently. So the rate of intellectual development is also a crucial factor and that, I think, might well be linked to the development of information processing. There is still a lot to find out about human intelligence and its development but, as these three books show, we already have some interesting information to consider. □

P.E. Bryant is Watts Professor in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Fruits of cognition

Stuart Sutherland

Cognitive Psychology and its Implications, 2nd Edn. By John R. Anderson. *W.H. Freeman:1985. Pp.472. \$25.95, £17.95.*

Cognitive Processes, 2nd Edn. By Lyle E. Bourne, Roger L. Dominowski, Elizabeth F. Loftus and Alice F. Healy. *Prentice-Hall:1985. Pp.404. \$28.95, £31.*

IN THE early 1960s, psychologists first began to notice the discoveries of workers in formal linguistics and artificial intelligence. Not the least remarkable thing about those discoveries was that they were made solely through the use of careful introspection and pure cerebration: never slow to learn from others, psychologists wisely decided to abandon behaviourism. Of course, psychologists being psychologists, they adhered to their dictum that nothing can be known unless a flawless experiment has been performed to demonstrate it, so they continued doing experiments. They replaced their studies of that fascinating beast, the laboratory rat, by experiments on mental processes: thus was cognitive psychology born. It has since flourished to such an extent that almost every other book in psychology has the word "cognitive" or "cognition" on the cover: indeed, even books on the psychology of women's liberation are nowadays ensured a sale only if they bear a title such as *Women and Cognition*.

Cognitive Psychology and Its Implications and *Cognitive Processes*, both revisions of successful texts first published five or six years ago, are part of the outcome of the cognitive revolution. They cover much the same ground — memory, attention, language, concept formation, problem solving and reasoning, though as to the last two topics it is somewhat unclear why if you are manipulating the tower of Hanoi, you are "problem solving", but if you are working on a syllogism, you are "reasoning". It is true that the latter problem could be solved by deduction, but in fact it rarely is except by logicians.

Anderson writes with more clarity and snap than do Bourne *et al.*, particularly when describing theories: his book could be read not only by students but by anyone who wants to know what is going on in the field. On the other hand, Bourne *et al.* describe the experimental evidence in much more detail and are not afraid to say that issues have not been resolved. Thus, in evaluating the Whorfian hypothesis (that language influences how the world is seen), Anderson claims the evidence does not support it, whereas Bourne *et al.* leave the question open.

Psychology is a maddening subject: it uncovers many issues but solves few. For example, both books take too sanguine a

view of the interpretation of Saul Sternberg's well-known experiment on how a person decides whether a test item belongs to a small set stored in immediate memory. Sternberg concluded that the test item is successively matched to each stored item and that, surprisingly, the search does not stop when the item is found but continues to the end of the list. This result was so exciting that a host of other experimenters promptly repeated the experiment with modifications cunningly devised to prove Sternberg wrong: whether basically right or not, his original interpretation is certainly very far from being the whole story. Part of the difficulty of securing firm explanations in cognitive psychology is of course that a person may simultaneously use several different

strategies in performing the same task.

Textbooks must simplify, and although both do so Anderson sometimes goes too far. Textbooks should also give the student some feel for the field, and here both seem a little wanting. Neither gives a clear account of the relations between psychology and linguistics or AI, and neither gives so much as a hint that most of the theories described are lacking in rigour. The books do not stand back from the hurly-burly of the journals to ask what work is likely to stand up in 20 years time and why. But within these limitations, which are the limitations of the subject, each in its own way does a highly competent job. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition, University of Sussex, Brighton BN1 9QG, UK.

Conversation in AI

John A. Campbell

Introduction to Artificial Intelligence. By Eugene Charniak and Drew McDermott. *Addison-Wesley: 1985. Pp.701. Hbk \$34.95; pbk £16.95.*

ARTIFICIAL intelligence (AI) has migrated only recently from postgraduate programmes (or from nowhere, in places that have upheld the view that AI is not intellectually respectable and even not quite nice) to undergraduate curricula in computer science. The migration has been hastened by demand, which is based partly on the observation that AI has already demonstrated its ability to be profitable. The problem which this change has highlighted is that AI overlaps so many other subjects that a first course in AI runs the risks of presenting knowledge in fragments and of having to begin with some remedial education.

No AI text which has been offered for undergraduate courses manages to avoid this problem. Those that come closest to success follow one of two routes: the axiomatic or the conversational. In the former, the primary emphasis is on the most far-reaching, basic or influential algorithms and operational methods. The good "conversational" texts include most of the same topics, but their first aim seems to be to compensate for students' gaps in background by presenting almost a transcript of a polished imaginary lecture series. A standard part of this approach is to state an idea or method, and then justify it by listing and explaining away the objections that may have occurred to the inventor or a not-unfriendly colleague.

Explanations of individual points by discussion in the conversational method may leave a student without the necessary background unmoved; the method has to rely on cumulative effect in introducing its

audience to the AI outlook. Nevertheless, some customers prefer this to the "axiomatic" approach, of which a good representative is *Artificial Intelligence* by P.H. Winston (Addison-Wesley, 1984). The present book is an equally good representative of the conversational school.

A further two-way division of general texts on AI is possible, between those that do and do not claim any relatively deep organizing principles for the subject. In this respect Winston again comes down on one side ("to make computers more useful") while Charniak and McDermott choose the other ("the study of mental faculties through the use of computational models"). Although this earns them extra points for accuracy — for example in their treatment of low-level processing of visual and some linguistic information — and for bravery, they do not succeed in turning their ambition into a backbone which unifies the lessons that a student may expect from a textbook. Such backbones are available: for example, the fact that (by approximation, good trial-and-error methods or ingenious representation of information) minds dispose of many kinds of problems without excessive computational effort while corresponding AI programs still contain no realistic means of escaping combinatorial explosions in numbers of paths to possible solutions. But this complaint applies to most AI textbooks, not just to the one under review.

Given the difficulties of presenting effective undergraduate courses in AI, this is a thorough survey of the whole field by experienced practitioners which is particularly effective on natural-language processing and (because of the limited competition) vision. People whose teaching or learning styles are more conversational than axiomatic may find it the most pleasing text to have appeared so far.

John A. Campbell is Professor of Computer Science at the University of Exeter, Exeter EX4 4PT, UK.

Case of a personal paradigm

C.R. Brand

Personality and Individual Differences: A Natural Science Approach. By Hans J. Eysenck and Michael W. Eysenck. Plenum: 1985. Pp.424. \$29.50, £25.65.

HANS Eysenck is Britain's best known and most cited psychologist. He is also the most reviled, not least for his ambitious plan (pursued now for more than 40 years) to bring the methods of natural and what he takes to be Kuhnian "normal" science to bear on human personality. For some time Eysenck has been a critic of the efforts of his indis disciplined peers to write textbooks of "personality theory" — books written according to a primitive calendar of saints that runs from Freud through the likes of Harry Stack Sullivan to "Ronnie" Laing, with barely a mention of the empirical efforts of Cattell, Guilford and himself. Here he is joined by his son in a courageous, readable and scholarly attempt to set the record straight.

It is a tribute to the Eysencks that even 358 pages of only occasionally self-indulgent text are barely able to encompass the many investigable contentions that Eysenck's theory has generated. As *cognoscenti* would expect, the Eysencks make the case that their four-and-a-half dimensions of personality — general intelligence (*g*), neuroticism (*N*), extraversion (*E*), psychoticism (*P*) and the "half" for lying (*L*) — account for a surprisingly high percentage of empirically registered variation in human behaviour and self-reported experience; and that these dimensions have demonstrable biological bases in genetic variations, in their physiological manifestations (for example in blood group), and in experimentally distinguishable psychological mechanisms — although Michael Eysenck sounds many cautionary notes about the latter possibility.

Altogether the book must appear an impressive achievement to those few humanistic or experimental (now "cognitive") psychologists who can be troubled to read it. Certainly one gasps at the amount of taxpayer's money that could have been better invested if psychological researchers had restricted their efforts of the past 20 years to the Eysencks' reasonable paradigm. Yet many will have their doubts — and not only about the startling disingenuity of the book's coverage of criminality, in the course of which the senior author (for it is surely he) interprets Passingham's sombre critique of the criminality-unconditionability link as "encouraging".

The general problem is that it is hard to

expand an originally three-dimensional model of personality (*g*, *N*, *E*) into which human variation was once ingeniously compressed without inviting the suspicion that perhaps a few more than four-and-a-half dimensions may finally prove necessary. Such a suspicion finds support in Cattell's typical recovery of at least six second-order dimensions and in similar recent findings in Baltimore by McCrae and Costa; but it becomes especially and even unduly compelling because of the authors' failure to delimit clearly what they sometimes call "the field" of personality to which their dimensions are meant to apply. Where, one asks of this putatively catholic theory, is any sustained representation of individual differences in motivations, attitudes, sensibilities, interests, moods and even abilities (apart from that mighty midget, little *g*)? There is some coverage of sex, Eysenck having learned the kind of thing that the empirically-minded public appreciates; but since we are invited to see the natural science beneath our pre-scientific habits of thought, it would be helpful to learn

something of how we are to translate into such discourse our commonplace descriptions of people as being variously endowed with properties of energy, conscience, affection and will — let alone of heart, mind, soul and spirit.

Unlike most psychological writing, this book addresses a large theme. It does so with mercifully infrequent use of words such as "sciamachy", "polythetic" and "intraception", and with a positive cornucopia (or "gallimaufry", as the Eysencks modestly prefer) of studies and experiments; and it contains many superb passages, for example on the "situationist" fad of social psychologists that we do not really have personalities. It will not, however, quite persuade anybody — at least when describing admired acquaintances — to abandon the language of the Bible and Shakespeare. That this language provides the Eysencks' only serious competition is, both fairly and unfairly, another measure of their success. □

C.R. Brand is a Lecturer in the Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK.

Head for education

Andy Young

Fundamentals of Human Neuropsychology, 2nd Edn. By Bryan Kolb and Ian Q. Whishaw. W.H. Freeman:1985. Pp.785. \$27.95, £24.95.

Mind and Brain: Principles of Neuropsychology. By Alberta Steinman Gilinsky. Praeger/Eurospan:1985. Pp.550. Hbk \$42.95, £47.25; pbk \$18.95, £21.50.

Brain and Mind. Edited by David A. Oakley. Methuen:1985. Pp.265. Hbk £15.95, \$36; pbk £6.95.

Two of these books (those by Gilinsky, and by Kolb and Whishaw) are general texts that cover a broad range of neuropsychological studies; the third, edited by Oakley, consists of a set of tutorial readings focused on a particular topic.

Kolb and Whishaw's *Fundamentals of Human Neuropsychology*, now appearing in a second edition, is one of the market leaders in its field. It is the kind of book that can convince sceptics that textbook teaching has a great deal to offer. It manages to be scholarly yet remain easily readable. If you use it to look up an unfamiliar topic, it succeeds in being informative, but gives both a clear impression that there is more to be learned and encouragement to the reader to follow things up. The authors thus manage elegantly to avoid the classic textbook failing of seeming to sum up a particular topic whilst actually only revealing a fraction of what is known. I can think of no better recommendation, and confidently expect the new edition

will be as popular as its predecessor.

I can't be so positive about *Mind and Brain*. This is a pity, because much of it is also well written and informative. The problem is that Gilinsky uses the book as a vehicle for developing a particular theory, the idea that "cognons" form the basic units of perception and cognition. In essence this is an (acknowledged) elaboration and modification of Konorski's notion of gnostic units. Cognons are claimed to account for a remarkable variety of phenomena, including letter and word recognition, face recognition, visual search, perceptual constancies, illusory contours, memory and conditioning.

All authors of textbooks must, of course, adopt some kind of theoretical stance, and it is perhaps commendable that Gilinsky's position is made so explicit. It is also unwise to dismiss new ideas just because they don't fit existing preconceptions. How we used to laugh at logogens. Nonetheless, I can't help but feel that a textbook is not the right place to discuss a theory which remains unproven.

In Oakley's collection of readings, a number of distinguished contributors discuss what is known about various aspects of the relation between brain and mind. The appearance of the book at this time is appropriate, because neuropsychologists now have the confidence to tackle such exciting and fundamental questions. This is an accessible anthology which students will find helps them get to grips with some of the most taxing neuropsychological questions. □

Andy Young is a Lecturer in the Department of Psychology, University of Lancaster, Lancaster LA1 4YF, UK.

Class of behaviour

Tim Roper

Social Behaviour in Mammals. By T. Poole. Blackie/Chapman & Hall: 1985. Pp.248. Pbk £13.95, \$22.50.

It is unusual nowadays to find a book on animal social behaviour that deals with a single taxonomic group or that arranges its subject matter on a taxonomic basis. The reason for this is simply that most recent work on social behaviour has revolved around functional questions, which often cross phylogenetic boundaries. A few taxonomic groups, such as the social insects, are specialized enough to be discussed as functional entities. But this is unlikely to be true of the class Mammalia. Can anything general but non-trivial be said about a group of species as diverse, in all respects, as shrews, elephants, bats, whales and humans?

The tension between ecology and phylogeny is evident in Trevor Poole's book, and is not successfully resolved. Poole begins with a brief summary of basic sociobiology and then deals in successive chapters with communication, reproduction, competition and resource exploitation. These are functional chapter headings, but the material within the chapters is often (but not always) arranged phylogenetically. Finally there is an order-by-order summary of social behaviour in mammals, starting with monotremes and ending with hominids.

The writing is clear enough but is factual rather than inspirational in tone, and has suffered from the need to compress so much information into a mere 200 pages. Because underlying principles tend to get lost in a mass of examples, I doubt whether the book will really inspire undergraduates. On the other hand, it is not a very satisfactory sourcebook either. Intricate lines of research and complex theoretical questions have had to be summarized in a sentence or two, and although in some places the references are surprisingly complete in others they are absent entirely.

Poole's book does remind the reader what enormous diversity there is within the mammals, and it corrects the common assumption that the only mammals whose social behaviour is at all interesting are the primates. It does not, however, convince me that mammalian social behaviour is a valid field in its own right, either for research or as the subject of an elementary textbook. □

Tim Roper is a Lecturer in the School of Biology, University of Sussex, Brighton BN1 9QG, UK.

• A review of four general textbooks on animal behaviour will appear in the *Nature* issue of 10 April.

Jolly impressive on primatology

J.R. Anderson

The Evolution of Primate Behavior, 2nd Edn. By Alison Jolly. Macmillan/Collier Macmillan:1985. Pp.526. \$24.95, £20.

ALTHOUGH it has remained useful and popular, the 1972 edition of Alison Jolly's textbook was showing its age. For some time, students have had to look elsewhere for information on developments in primate ecology, sociobiology and psychology. The long-awaited second edition of *The Evolution of Primate Behavior* now makes life easier. Like its precursor, the book provides a scholarly and wide-ranging survey of current research in primate behaviour, in an informal style that will be widely appreciated.

Comparisons with the first edition are instructive. Certain earlier conclusions have not been borne out (for example Jolly acknowledges that the data now available do not indicate any strong relationship between folivory and territoriality in primates), while other questions raised in 1972 have been answered positively: thus we now know that non-human primates indeed show a similar progression through the stages of sensorimotor intelligence to that seen in human beings. Elsewhere, some topics that were in fashion in 1972 now get less space. Species differences in aggression, and environmental influences on aggression drop out (although competition between individuals and groups remains an important theme); the differen-

tial rearing experiments of the 1960s are described quite briefly; and "attention structure" doesn't even make it into the new index. These will not be badly missed, given the coverage of topics of current interest that were virtually non-existent in primatology 15 years ago, such as the influence of plant secondary compounds on feeding, social awareness and deception, kin recognition, reciprocal altruism and adjustment of sex ratios of offspring.

As in the first edition, the chapters on intelligence are among the strongest. Here again there are some changes of emphasis in the author's summaries of the recent literature on tool-use, linguistic abilities and "cognition". Developments in ape language research are described at length, as are Piagetian approaches to cognition. Correspondingly, the treatment of other laboratory techniques seems brief or is missing (for example cross-modal matching, delayed response and its variants, serial position functions).

One unfortunate feature of the new edition is the many bibliographical errors: over 50 references cited in the text are missing from the bibliography, and there is a remarkable number of mismatches between dates in the text and bibliography. However, this will not unduly worry undergraduates looking for an up-to-date and readable account of the field. There are better recent books on some of the subjects discussed by Jolly, but, as a single textbook covering the spectrum of modern behavioural primatology, this is the best there is. □

J.R. Anderson is Maître-Assistant Associé in the Laboratoire de Psychophysiologie, Université Louis Pasteur, 7 rue de l'Université, 67000 Strasbourg, France.

Underlying action

Felicity A. Huntingford

Essentials of Behaviour Genetics. By David A. Hay. Blackwell Scientific:1985. Pp.359. Pbk £15.80, \$27.

Essentials of Behaviour Genetics is an authoritative and valuable book. The introductory and concluding chapters provide interesting accounts of the history and probable future directions of the subject, but the book's particular strength is in its careful discussion of techniques for identifying genetic influences on behaviour and the mechanisms by which such influences act. These are illustrated with reference to different genetic mechanisms (from single locus mutants and detectable chromosome anomalies to polygenic systems) and different aspects of behaviour (from courtship and habitat selection to learning and personality traits).

Many different aspects of behaviour genetics are explored, including its relevance to other disciplines such as psychology and sociobiology, and its implications for our understanding of evolution. Hay also gives informed consideration to the social implications of the subject, in particular those of research into the inheritance of IQ. His extensive use of quotations helps to give a balanced picture of a highly controversial topic.

The book is not without its weaknesses. In particular, I question whether it is entirely suitable for its intended audience, which is defined in the preface as advanced students in psychology and genetics. Recognizing the disparate needs of these two groups, Hay introduces the psychologists to genetics by discussing the inheritance of mental retardation. For students whose background is in modern genetics, behaviour is treated as a particularly challenging phenotype for genetic research.

Neither of these good ideas quite works out in practice. The introductory chapter

contains a useful section on the need for adequate behavioural measures ("garbage in — garbage out"), but many of the examples use behavioural measures which are anything but adequate. The limitations are discussed, but students of genetics may well go away with the (false) impression that much behavioural research is of rather poor quality. Basic genetic concepts are introduced in Chapter 2 (on mental retardation, and containing gruesomely fascinating examples of congenital disorders), but without much explanation. Interesting as it is, the chapter does not really provide the clear statement of the underlying principles that a newcomer needs.

The degree of detail with which the subject is treated is inconsistent and, in places, excessive; important concepts are sometimes obscured by too many equations and figures. For example, in Chapter 3 on polygenic inheritance, is it really necessary to show how one calculates a correlation coefficient? Hay is right to stress the importance of methodology, but the reader of a book on the *essentials* of behaviour genetics doesn't need to be able to apply the techniques, only to understand the principles involved.

The sections giving details of methodology make the book rather long, and although the pages bear the hallmark of Blackwell's good design team the print is comfortably cramped in places. Perhaps to keep the book within reasonable limits, Hay concentrates strongly on work on human beings and rodents, although, as he points out himself, behaviour geneticists are turning increasingly to different animal groups. Other vertebrates get no more than a passing mention and, while there is an interesting chapter on invertebrates, this is relatively brief. Here is one place where more detail would have been welcome.

Finally, I have to mention the cartoons, a number of which are included to enliven the text. Some of them (for example, Frank Beach's picture of the white rat of Hamelin piping a crowd of smiling scientists, all clutching the tools of their trade, towards the banks of a torrential river) are both witty and pertinent. However, most of them seem to serve no useful purpose; crudely drawn pictures of mice in bikinis and flies grasping bunches of flowers did not help this particular reader through the text.

Existing textbooks on behaviour genetics, although in some respects not so impressive as Hay's, make the topic more accessible to the student with a limited background in genetics. But as a detailed and wide-ranging review for the more specialist reader, this book has a great deal going for it. □

Felicity A. Huntingford is a Lecturer in the Department of Zoology, University of Glasgow, Glasgow G12 8QQ, UK.

Nervous beginnings

Maria Fitzgerald

Principles of Neural Science, 2nd Edn. By Eric R. Kandel and James H. Schwartz. Elsevier: 1985. Pp.979. Dfl.160, \$47.50.

Progress in Neuroscience: Readings from Scientific American. Introductions by Richard F. Thompson. W.H. Freeman: 1985. Pp.151. Hbk \$21.95, £21.95; pbk \$12.95, £12.95.

The Brain: An Introduction to Neuroscience. By Richard F. Thompson. W.H. Freeman: 1985. Pp.363. Hbk \$29.95, £33.50; pbk \$17.95, £17.95.

Fundamental Neuroanatomy. By Walle J.H. Nauta and Michael Feirtag. W.H. Freeman: 1985. Pp.340. Hbk \$39.95, £39.95; pbk \$26.95, £18.95.

THERE is nothing so welcome to student and teacher alike as a new edition of a well-loved textbook. In the case of Kandel and Schwartz's *Principles of Neuroscience*, the publication of a second edition will also be good news for research scientists. When the book first appeared in 1981, it was immediately recognized, despite its idiosyncrasies, as the kind of wide-ranging, high-quality text that was badly needed in both the classroom and the laboratory. The new edition is not only updated, it is improved and somewhat enlarged with ten more chapters and a fully comprehensive index.

The two principal trends in neuroscience of recent years, increased study of the molecular mechanisms underlying neural transmission and of the development of the nervous system, are now given considerable attention. The greatest changes, however, reflect the main interests of the authors and their contributing colleagues, cell biology and behaviour. The opening parts of the book on the cellular and molecular biology of the neurone are therefore especially good, encompassing cytology, synthesis and distribution of neuronal proteins, axon transport and the cytoskeleton, with an enlarged section on glia. The account of membrane biophysics has been updated to include molecular characterization of ion channels, and the chapters on synaptic transmission now introduce patch clamping, M-channels, slow synaptic changes, peptides and events at the molecular level. The sections dealing with higher cortical functions and behaviour are of a similarly high standard. Behaviour is discussed in terms of genetic and environmental influences in the whole animal and at the cellular level, from its emergence in the embryo to changes associated with old age.

The central chapters deal with what might be termed "functional" neurophysiology, that is, sensory systems, motor systems, brain stem and reticular forma-

tion, hypothalamus and limbic system. This is the weakest part of the book, because the authors make the mistake of dissecting sensory and motor systems down into labelled lines, hierarchical relationships, maps, specific pathways and fixed receptive fields; it is as if everything is wired together to perform a specific function and does so only in parallel with other such wires. The mammalian CNS is treated as a series of marine snail neurones. Such an "apysiological's" approach provides only a limited view of the organization of the mammalian CNS, how it interacts with the outside world and changes with experience. Pain, for example, a very human experience, is particularly poorly covered.

Yet it is hard to criticize a book which has so much to commend it. Of particular importance is that there is throughout a strong emphasis on brain disorders and neurological syndromes, examples of which are used in conjunction with details of neurophysiology, pharmacology and anatomy to explain brain function. It is this clinical emphasis, exemplified by appendices on brain fluids and their disorders, cerebral blood flow and stroke, along with the coverage of neural mechanisms, that puts *Principles of Neuroscience* in a class of its own.

Another excellent read is *Progress in Neuroscience*, a collection of articles from *Scientific American* published between 1979 and 1984. All of them are by leading scientists and together they make a useful and approachable overview of recent advances in the field. There are, for example, chapters by Llinas on calcium at the synapse, Iversen on neurochemistry, Cowan on neural development, Knudsen on hearing, and Scheller and Axel on genetic control of behaviour. Each of them is written in a clear and accessible style, and there are lots of coloured diagrams. Inevitably, some of the articles are a bit out of date now, but this does not matter too much because by pin-pointing fast-growing areas and conveying the excitement of new discoveries the reader is encouraged to search out more recent publications.

The editor of this anthology, R.F. Thompson, suggests that the readings can also complement his new book, *The Brain: An Introduction to Neuroscience*. This may be so to a certain extent, but his text is really for a rather more limited audience. It aims to cover the workings of the brain and its nerve cells for readers with no background in science, and does so admirably. This is the first book of this kind to deal with all the basic topics in neuroscience from membrane potentials to anxiety neuroses at an introductory level. As such it will be ideal for those such as speech scientists or human biologists studying neuroscience for the first time.

As well as chapters on membrane and

synaptic physiology, neurotransmitters, and sensory and motor processes, there is also good coverage of peptides and hormones, development and ageing, and finally an introduction to studies on behaviour. The emphasis is on all vertebrates, not just human beings. The whisker barrels of the rat and the hearing apparatus of the barn owl are given as much prominence as the equivalent sensory apparatus in man, and each is described in relation to the animal's needs. It is a pity that many of the key references at the end of each chapter are rather out of date, because this is certainly not true of the ideas presented in the text.

When an editor of *Scientific American* has got together with a leading neuroanatomist to produce a new textbook, one might expect something really special. In the case of Nauta and Feirtag's *Fundamental Neuroanatomy*, however, the end result, although interesting, is not a complete success. The book is smartly presented with large margins, red chapter headings and lots of illustrations, but it has many faults as a textbook.

Like all good teachers of neuroanatomy, Nauta is clearly attempting to explain the overall organization of the brain as well as its detailed connections. To that end the book is arranged such that after an introduction on techniques and phylogenetic aspects, there are sections on sensory and motor pathways, the autonomic nervous system, and cortical and limbic functions, before the account of core neuroanatomy. The problem is that the journalistic input, which could have made the subject more approachable, has resulted in a irritating and self-consciously written book. The arrogant indifference to punctuation and chatty comments, one after the other, make it almost impossible to pick out the facts. Some readers may love the confident "let me take you through the brain" style, others may be amused by it but few will learn from it. The section on connectivities is particularly bad, with a series of horribly complicated brainstem diagrams, in which, on top of everything else, the cell bodies are drawn as arrow heads and the terminals as dots so at a glance everything seems the wrong way round.

Nevertheless, when the neuroanatomist is allowed to dominate the better aspects of the book come through. The strongest part is the atlas of cross-sections through the brain and the accompanying text which, when all is said and done, is one of the best ways of learning neuroanatomy. I would suggest reading this book having already learnt the fundamental facts, perhaps during revision. It has one important strength — it may be irritating but it certainly isn't boring. □

Maria Fitzgerald is a Lecturer in the Department of Anatomy and Embryology, University College London, Gower Street, London WC1E 6BT, UK.

Beyond the facts of pharmacology

Donald Jenkinson

Goodman and Gilman's The Pharmacological Basis of Therapeutics, 7th Edn. Edited by A. Goodman Gilman, L.S. Goodman, T.W. Rall and F. Murad. Macmillan/Collier Macmillan: 1985. Pp. 1,872. \$65, £65.

Lecture Notes on Pharmacology. By H.F. Grundy. Blackwell Scientific: 1985. Pp. 446. Pbk £9.80, \$12.95.

Learning Pharmacology Through MCQ: A Comprehensive Text. By B.J. Large and I.E. Hughes. Wiley: 1985. Pp. 204. Pbk \$11.10, £6.95.

Gaddum's Pharmacology, 9th Edn. By A.S.V. Burgen and J.F.M. Mitchell. Oxford University Press: 1985. Pp. 260. Pbk £9.50, \$17.95.

FOR many years most textbooks of pharmacology were by necessity little more than compendia of factual information, much like early texts on organic or inorganic chemistry. In due course, the latter were to be transformed as mechanisms and the basis of reactivity became understood. A similar development is taking place in pharmacology as the molecular basis of drug action is unravelled, and as the first receptors are isolated and their structure and exact function determined. At the same time, a great deal has become known about the absorption, distribution and fate of drugs in the body (pharmacokinetics). This topic, along with bioassay, drug development and the study of receptors, is described as general pharmacology, and the writer of a pharmacology textbook has now to decide how, and to what extent, to incorporate the new material. Not surprisingly, these four books deal with the problem in a variety of ways.

Goodman and Gilman is indisputably the most important source book on pharmacology as related to therapeutics, and the publication of the seventh edition will be greatly appreciated by students and their teachers alike. The book has been thoroughly revised, and in some areas expanded, as in the useful and carefully documented 65-page table of pharmacokinetic data which first appeared in the preceding edition.

It has to be said, however, that some of the new material in the first chapters on general pharmacology is not yet of the same high standard as the rest of the book. Though Chapter 2, on the mechanisms of drug action, and concentration-effect relationships, begins excellently, and is first-rate on the biochemical consequences of receptor activation, it is quite unsatisfactory to commence a detailed discussion of the quantitation of drug action in terms of

the Michaelis-Menten equation and the Lineweaver-Burk plot. This approach provides a far less suitable framework for the discussion of dose-response curves and competitive antagonism than the long-established treatments by J.H. Gaddum, H.O. Schild and R.F. Furchgott.

This section also ends unhappily with an account of partial agonists which fails to distinguish between intrinsic activity and efficacy. For 30 years, pharmacologists dealing with receptors have used the term efficacy in the specific sense introduced by R.P. Stephenson, to whom we owe the convention that a partial agonist which can produce at most a half-maximal response in a tissue has an efficacy of unity. It is disconcerting then to see on p. 43 of Goodman and Gilman a full agonist categorized as having an efficacy of 1! The distinction might seem pedantic, but there are now many examples to show that careful thinking about the expression and interpretation of partial agonism can throw valuable light on the classification of drugs and receptors. It would be unfair to end on such a negative note, however. In general all concerned have done an excellent job on the revised edition of this classic work.

Teachers of pharmacology to medical students have been looking forward to a new *Lecture Notes on Pharmacology*, to succeed the good but now dated volume by J.H. Burn. Preclinical students found this invaluable, not least for last-minute revision. How, then, does the new volume compare? Though it contains a great deal of excellent material (e.g. on cardiovascular pharmacology), overall the book is a little disappointing. For one thing there is a marked imbalance in the coverage; some topics (the adrenergic neurone, the mechanism of action of local anaesthetics) are considered in much greater detail than others (drugs affecting the kidney). There is also a rather uneasy blend of basic with very recent and advanced material, some of which (inevitably, perhaps) is incorrect. For example, agonists and antagonists are said on p.92 to combine with different subunits of the acetylcholine receptor. This is not in keeping with the description elsewhere in the book of antagonists acting on the nicotinic receptor at the neuromuscular junction as competitive, or indeed with recent evidence. Cause and effect are reversed in a passage on receptor-mediated ion conductance changes on p.100, and the extracellular role ascribed to cyclic AMP on p.109 is surprising. These, however, are matters which can easily be corrected or clarified in a new edition, when the opportunity could also be taken to improve some of the illustrations.

Anyone who has tried to set MCQs, and to assess those of others, will know that only one's own seem to be free from ambiguity (until, that is, they are seen by some-

one else, preferably from another institution). *Learning Pharmacology Through MCQ* is attractively set-out and includes almost 200 wide-ranging and generally good multiple choice questions, together with answers and concise explanations, but it contains a number of unfortunate errors which should have been caught at some stage.

Question 67, for example, is marred by the substitution of an α for a β in the opening statement, and students will be puzzled by item 4 of Question 22, where the statement that the affinity constant of an antagonist will be a larger numerical value the more potent the antagonist is said to be false. It is also inappropriate to include proprietary with non-proprietary names, as in the trick completion to Question 119, item 1, where the reader is asked to compare the hypnotic powers of nitrazepam and Mogadon! A corrected and expanded edition would be valuable, though my own experience has been that

MCQs are likely to be of more help in teaching than in learning.

Finally, Gaddum. This book has long been the standard introduction to pharmacology and the comprehensively revised new edition will continue to serve this role. The authors have managed to introduce much fresh material (for example on peptide neurotransmitters, interestingly renamed as neuromessengers), while at the same time retaining the essence of Gaddum's succinct, balanced style, and his emphasis on principles and experimental evidence. This book will form an excellent complement to more detailed texts such as Katzung's *Basic & Clinical Pharmacology*, a second edition of which appeared in 1984, and the forthcoming successor to Schild's *Applied Pharmacology*, now in press. □

Donald Jenkinson is a Professor in the Department of Pharmacology, University College London, Gower Street, London WC1E 6BT, UK.

On cancer's track

I. Bernard Weinstein

The Molecular Basis of Cancer. Edited by Peter B. Farmer and John M. Walker. Croom Helm/Wiley: 1985. Pp.349. Hbk £25; pbk £12.95, \$29.95.

RUDOLF Virchow, the father of cellular pathology, stated that no one, even under threat of torture, could define cancer. Approximately one hundred years later we are, at last, approaching a definition of the disease at both the cellular and molecular levels. These developments provide the subject matter for *The Molecular Basis of Cancer*, a text designed for students entering cancer research who are already conversant with basic biochemistry. The book is quite comprehensive, and spans the range from clinical and pathological features of the disease to biochemical details of cancer causation and therapy. As a collection of chapters by six authors, however, it suffers from repetition and lack of overall coherence.

Cancer research is moving so rapidly that although there is a detailed description of the cell surface, exciting new information on receptors, growth factors, protein kinases and signal transduction is not covered. The reader will also have to look elsewhere to learn of the intriguing roles that certain oncogenes play in controlling the transcription of specific genes. The plethora of oncogene changes now revealed in rodent and human tumours also supersedes much of the information given here. Furthermore, the pre-eminent role often assigned to dominant acting oncogenes must be reconsidered, in view of the fact that in certain hereditary can-

cers (retinoblastoma and Wilm's tumours) the loss of function of recessive genes (perhaps anti-oncogenes) plays a role in tumour formation. Normal genes that can suppress the tumour phenotype are also being revealed in cell hybrid studies.

Absent too is the new information on human tumour viruses, including the causative role of human papilloma viruses in cervical cancer and HTLV II in hairy cell leukaemia. Similarly, although the accounts of chemical and radiation carcinogenesis and short term bioassays are highly informative as far as they go, they do not discuss the roles of activated forms of oxygen in cancer causation, the role of protein kinase C and phosphoinositol turnover in tumour promotion, the development of highly sensitive immunoassays and post-labelling methods to detect carcinogen-DNA adducts, and the opportunities offered by "molecular epidemiology" for defining more precisely the roles of environmental factors in the aetiology of human cancers.

Very recent advances related to diagnosis and therapy are also not included, but it would be churlish to continue with a catalogue of omissions from the book. Rather, the above comments should be taken not so much as criticism of the various authors but to indicate the impossibility of compiling a fully up-to-date account of cancer research. As it stands, the book is a valuable primer for students entering the field. But they will have to read the contemporary literature assiduously to track the lightning pace at which cancer is being defined in molecular terms. □

I. Bernard Weinstein is Director of the Columbia University Cancer Center and Frode Jensen Professor of Medicine, Health Sciences Center Building, 701 West 168th Street, New York, New York 10032, USA.

Old and Primrose come of age

Stephen Oliver

Principles of Gene Manipulation: An Introduction to Genetic Engineering, 3rd Edn. By R.W. Old and S.B. Primrose. Blackwell Scientific: 1985. Pp.409. Pbk £11.80, \$21.

IT HAS become commonplace for organizers of symposia on molecular biology to open proceedings by showing a slide to illustrate the exponential increase in the thickness of the abstracts books through successive meetings. Something similar is happening to Old and Primrose's *Principles of Gene Manipulation*; the third edition is twice the size of the second. However, this standard text has not merely put on weight it has also, like the subject it describes, gained in maturity. This shows itself in a greater sureness of touch in the opening chapter, where the basic techniques are laid out, and also in a greater willingness to discuss the problems and pitfalls in using gene technology and how these may be solved or circumvented.

New chapters have been added on bacteria other than *Escherichia coli*, and on yeast and other microbial eukaryotes; on microinjection of DNA; and on industrial applications (including patent law). Other chapters, such as those on cloning in plant and animal cells, are much expanded. Ironically, but inevitably, some of the newest sections seem the most dated. This is particularly true of the account of site-directed mutagenesis, an area in which there has been rapid progress since the end of 1983 when the literature search for this volume appears to have been completed. There are also some surprising omissions. Cloning in *Streptomyces* is dismissed as being too similar to *E. coli* and, in any case, reviewed elsewhere. This biologically distinctive and industrially important group of organisms surely deserves more extensive treatment in the fourth edition.

The increased size of the volume has permitted the authors more space to describe the biology of the systems they deal with. This makes for stimulating reading which places the technology in context. In the chapters on applications, it is stressed that the main importance of recombinant DNA technology has been to extend and deepen our knowledge of molecular biology. This third edition will do the same for a new generation of students, who will also appreciate its reasonable price tag and clear, informative diagrams. □

Stephen Oliver is Senior Lecturer in Applied Molecular Biology at the University of Manchester Institute of Science and Technology, PO Box 88, Manchester M60 1QD, UK.

A macromolecular build-up

C.D. Rodger

Molecular Structure: Macromolecules in Three Dimensions. By Robert J. Flet-terick, Trina Schroer and Raymond J. Matela. *Blackwell Scientific*:1985. Pp.196. Pbk £19.80, \$19.95.

THE greater availability these days of molecular model-building kits has encouraged many of us teaching undergraduates in biology and biochemistry to supplement the theoretical material with practical "molecular construction" exercises. *Molecular Structure* is aimed at those using this approach. The book is unusual, however, in that it is attempting to be two things in one — a handbook for use with the molecular model-building kit also produced by Blackwell, and an explanatory textbook in its own right. Book and modelling kit can be purchased separately, though it is difficult to review the book without including some reference to the models as well.

The text is divided into three parts. The first ("Molecules and Modelling") contains a general introduction to molecular symmetry and the Blackwell modelling

system. The Blackwell models are not of the "all-atom" type and are intended for the construction of the backbones of proteins and nucleic acids, but not the side chains. They enable particularly accurate setting of the backbone angles because a small vernier is built into each of the components, thus allowing the angles to be set to an accuracy of 1°. Despite this precision, it seems that all the exercises in the book can be accomplished in a single practical session.

Part II, "Protein and Nucleic Acid Structures", introduces the concepts of folding and domains in polypeptides, and then continues with more detailed accounts of the secondary structures of polypeptides, together with instructions for constructing them from the kits. The same format is then used for describing nucleic acid structures. The descriptions are quite thorough, but the resulting backbone models of secondary structures seem rather crude by comparison. I suspect many would not consider the time spent building these particular structures as backbone models to be worthwhile at undergraduate level; such models cannot demonstrate useful concepts, such as the dominance of side chains in the α -helix or the relative tilts of bases in different nucleic acids, as well as all-atom kits. The backbone kits come into their own, however, when the book explains the construction of transfer RNA at the end of this section; definitely not a task for the all-atom approach in a student practical!

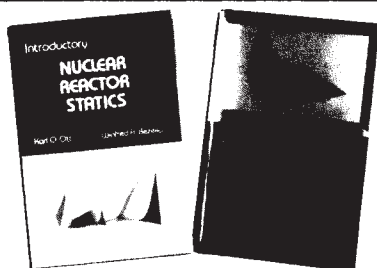
Part III contains instructions for building more complex structures such as insulin, the *cro* repressor and lysozyme. Although these are more complicated molecules, construction time is still within three hours because they are backbone models. It is with these structures that the backbone kits are likely to be most useful.

At first glance it appears that saccharide conformations are ignored altogether. However some details are included in Part III, in the instructions for building the lysozyme substrate containing *N*-acetylmuramic acid and *N*-acetylglucosamine. The authors have recognized that the backbone type of molecular modelling is unsuitable for demonstrating the intimacy of active-site substrate complexes and they include guidance on the use of Lab-Quip model kits for this exercise.

The text is well set out and easy to read. It contains many illustrations, albeit with rather sparing use of colour, and references are included at the end of each chapter. Representing three-dimensional structures on the flat page is not easy, and I suspect that some of the illustrations may not in fact help the reader to visualize the three-dimensional arrangements as the authors intended. Nevertheless, in its role as a companion handbook to a model-building kit the book should prove very useful. □

C.D. Rodger is Biology Course Leader in the Department of Sciences, Roehampton Institute of Higher Education, London SW15 5PJ, UK.

NEW TITLES BY ANS



"Introductory Nuclear Reactor Statics"
by Karl O. Ott
Winfred A. Bezella

Recommended as an undergraduate or graduate textbook. Valuable source book for reactor analysis and design.

"Introductory Nuclear Reactor Dynamics"
by Karl O. Ott
Robert J. Neuhold

Examines fundamental principles of Nuclear Dynamics. Companion volume to Nuclear Reactor Statics, published in 1983.



For additional information write:

American Nuclear Society
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Biology's conscience

E.G. Richards

Physical Chemistry: Principles and Applications in Biological Sciences, 2nd Edn. By Ignacio Tinoco, Jr, Kenneth Sauer and James C. Wang. *Prentice-Hall*:1985. Pp.706. \$38.95, £41.70.

PHYSICAL chemistry, it has been said, is the conscience of the biologist. We expect the mechanisms propounded by biochemists and molecular biologists to conform to the laws of physics, and because we are dealing with molecules, physical chemistry is the bridge. Even if one does not take this reductionist position, it must be admitted that the subject is at least the hand-maiden of the modern biology which would be infinitely poorer were it not for such experimental techniques as ultracentrifugation, electrophoresis or X-ray crystallography.

In the new edition of their book, Tinoco, Sauer and Wang, all expert practitioners of physical chemistry, provide the foundation for a sound understanding of the subject, as well as insights into some of the important techniques. The book is

specially written for the biologist, by which the authors mean the biochemist, biophysicist and molecular biologist; as such it invites comparison with Cantor and Shimmel's three-volume *Biophysical Chemistry*, but it is more appropriate for the beginner and thus fills an important gap. The coverage of fundamentals is good and geared to the interests of the intended audience. Some of the more esoteric methods, however, such as inelastic light scattering, are discussed but briefly with mere quotation of the relevant formulae.

One invaluable feature is the large number of worked examples in the text and of problems for the student (about one-third with answers at the back). Another, at any rate of the first half of the book, is the absence of higher mathematics, though more mathematical expertise is required for the later chapters on X-ray diffraction and quantum mechanics. The second edition has been improved by the near universal use of SI units, as well as the inclusion of extra topics. Altogether, it should remain a very popular textbook. □

E.G. Richards is a part-time lecturer in the Department of Biophysics, Cell and Molecular Biology, King's College London (KQC), 26-29 Drury Lane, London WC2B 5RL, UK.

Lewin renewed

W.J. Brammar

Genes II. By Benjamin Lewin. Wiley: 1985. Pp.716. Hbk \$52.55, £40.35; pbk \$20.70, £16.75.

THE first edition of *Genes* was well received as a balanced overview of an important and rapidly developing subject. The field has moved a long way in the two years since the first edition and Dr Lewin has now revised the book in the attempt to stay "au courant with present opinion".

In his preface the author apologizes for the fact that "the brute force of DNA technology has . . . obscured the elegance of traditional genetics". As editor of *Cell*, he is in a strong position both to report and to influence the balance between brute force and elegance in molecular genetics.

The main discernible change between the first and second editions is in Lewin's perspective on the subject matter. The introductory chapters have been reworked, with more emphasis on the biochemistry of heredity and the structure of DNA. The replacement of "What is a Gene? A Genetic View", the first chapter, by "DNA is the Genetic Material" gives the sense of the change in approach. There are times when the biochemistry is less than authoritative, most seriously perhaps when the definition of the crucial hydrogen bond excludes two of the five types of such bond that are responsible for the duplicity of the DNA helix. *Genes II* should not be judged as a biochemistry text, however, and it does provide a readable, well-illustrated and detailed account of the most important aspects of molecular genetics.

By judicious rewriting, Lewin has made space for new material without increasing the overall length of the text. Many developments at the forefront of the subject, including the structure of the T-cell receptor, homeotic sequences and the powerful transgenic animal technology are presented, and a new final chapter discusses ways of "engineering changes in the genome" to study aspects of tissue-specificity and the control of gene expression.

Like its predecessor, this edition of *Genes* can be highly recommended as a text suitable for advanced undergraduate teaching of molecular genetics. The carefully selected references to important reviews and research papers at the end of each chapter will be invaluable in guiding the serious student more deeply into the subject. But make haste: even while you read, the frontiers are advancing and Dr Lewin is doubtless at work on *Genes III*. □

W.J. Brammar is a Professor in the Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK, and Director of the University/ICI Joint Laboratory for Molecular Genetics.

Figuring out what's happening in immunology

William E. Paul

Immunology.* By Ivan Roitt, Jonathan Brostoff and David Male. Churchill Livingstone/Mosby: 1985. Pp. 320. Pbk £12.95; hbk \$21.95.

Introducing Immunology. By Norman Staines, Jonathan Brostoff and Keith James. Gower Medical, London:1985. Pp.59. Pbk £1.95.

THE rise of immunology as a biomedical science has been matched by the increasing recognition of the central role of the immune system in health and disease. This has underlined the need for texts designed for undergraduates, graduates and students of medicine, and for books which are aimed at an intelligent non-professional audience. *Immunology* is intended as a first text in immunology accessible to undergraduates but with sufficient content to be of value to starting graduate and medical students. *Introducing Immunology* explains the immune system to a general audience in a very concise format.

Immunology is a remarkable book which sets a new standard for introductory texts on the subject. It covers authoritatively the basic areas of immunology and deals, in an introductory way, with many of the key aspects of clinical immunology. But its most innovative and striking feature is the extensive use of figures, many of which are very handsome. For example, Chapter 1 ("Adaptive and Innate Immunity") consists of nine pages which include 25 figures, many of them in four or more colours. These illustrations were obviously planned with care and have been executed with considerable style; indeed, study of the figures and legends alone would serve as a fine introduction to immunology.

Immunology places heavy reliance on the illustrations in an effort to create a superior teaching device. To a large degree this succeeds, although I sometimes found them distracting as they often break the logical chain of development of specific issues. Furthermore, in many cases there is considerable redundancy between the text and the figure legends, so that there are essentially two descriptions side-by-side of the same material, both in a highly telegraphic style. Indeed, some of the fig-

**Slide Atlas of Immunology*, which contains 750 transparencies of illustrations in *Immunology*, together with supporting material, is available from Gower Medical in London and New York, and C.V. Mosby in St Louis, Missouri, price £500, \$850.

ures illustrate points which could have been made in the text alone, and elimination of them would have allowed a better presentation of the main points in each chapter.

The authors received a great deal of expert assistance in the preparation of the book, as noted in their acknowledgements. This has lent authority to each chapter, while the overall direction of Roitt and his colleagues has maintained continuity and uniformity of style. However, while obvious efforts have been made to keep the book up to date, it appears that the detailed planning and the time necessary to produce this unusual volume may have led to a relatively long period of writing and production. Some chapters, then, are a bit dated, though the passages concerned can be put right in the second edition which will undoubtedly follow.

It would be unfair to close this review on even a modestly negative note. *Immunology* breaks new ground and all introductory texts on the subject will in future be judged against it.

Introducing Immunology has a very different purpose. Its goal is to provide a short picture of immunology for secondary school and university students, and for the general public. It represents an effort, supported by the British Society for Immunology, to encourage students to consider a career in immunology and makes some practical suggestions concerning opportunities for education and degrees in immunology in Britain.

In less than 50 pages, the authors present the major elements of the immune system and discuss aspects of clinical immunology. Again, the figures are outstanding, which is no surprise because the production team was the same as for *Immunology*.

This slender volume does not stint in its presentation of the technical aspects of the field. A good example of its depth is found in the section on complement, where the principal components of the classical and alternative pathways are presented in a rather sophisticated manner. For both the general reader and the student seeking a brief introduction to immunology, *Introducing Immunology* will be a challenging but rewarding experience. □

William E. Paul is Chief of the Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA.

Textbooks supplement — prices

Where possible both dollar prices in the United States and sterling prices in Britain are given in the bibliographical details for each book. Export prices will generally be higher. If a dollar or sterling price is not cited, the book is not available in one of the two countries. However readers should check both price and availability of books before ordering.

Down to development

Michael Akam

From Gene to Animal: An Introduction to the Molecular Biology of Animal Development. By David De Pomerai. Cambridge University Press: 1985. Pp. 293. Hbk £25, \$44.50; pbk £8.95, \$14.95.

A SHORT text on the molecular biology of development must be a masterpiece of selection. Ideally, such a work would define questions in the language of embryology, and answer them in the language of molecular biology. We would like to know what molecules constitute maternal information, how interactions between these and the zygote genome generate more complex patterns, what description of cell types and morphology is appropriate to match the internal molecular logic, and how the differentiation of each particular cell type is executed through hierarchies or networks of control. Although such a work cannot yet be written, we might hope that a textbook for today's undergraduates would at least be able to set the scene, show where real progress has been made and, above all, convey the promise of the near future.

From Gene to Animal does not wholly succeed in meeting this agenda, in part because of the way that the book is organized. De Pomerai separates an elementary but extensive discussion of eukaryote molecular biology (the first half of the book) from a consideration of the molecular strategies of development which are condensed into one central chapter. As a result the significance of molecular mechanisms for developmental regulation sometimes slips between the cross references.

When molecular techniques are applied to the study of development, the first questions to be answered may well be the last things a student needs to know (the physical map of a gene is not always its most pertinent characteristic). De Pomerai does not go far enough to redress this bias. Much of the detail which he includes does little to illuminate the mechanism of development; the organization of histone and tRNA gene families tells us more about genome evolution than developmental regulation. Elsewhere five pages on nucleosomes could have been condensed to one longish paragraph without losing the main points.

The final three chapters are devoted to case studies. Like the rest of the book they are well written and nicely presented. They contain much useful information, and most of it is accurate, so far as I can see. But the accounts of erythroid differentiation and of the oviduct epitomize, unconsciously, the shortcomings of much molecular developmental biology. A

short introductory section in each chapter presents a classical, almost histological description of developmental events. The story then jumps to molecular details of terminal gene expression. In between lie many of the salient problems of development, but apart from passing reference to hormonal effects these are not considered. It would have been nice to read at least some discussion of how cellular events in erythropoiesis can be investigated, for example by the use of erythropoietic cell lines.

The final chapter covers insect development; unfortunately, it was written at just the wrong time. Now, I would choose insects (more precisely, *Drosophila*) to illustrate a number of key topics. One would be the significance of germ-line transformation for studies of gene control elements, and for the manipulation of development by the engineered expression of regulatory genes. And while P-element-mediated gene transfer does receive mention, it is only in the extraordinary context of chromosome puffing. This passing reference will hardly guide a student's perception of the field, and can only be explained by its hurried insertion into a near-final text.

Similar considerations no doubt excuse the short shrift given to recent demonstration of patterned gene expression in the

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

J.E. Sulston

The "late pretzel" stage of embryogenesis in the nematode *Caenorhabditis elegans*. The picture is taken from *Genetic Analysis of Animal Development*, by Adam S. Wilkins, published by Wiley. Price is \$69.95, £71.55.

Drosophila blastoderm, although these results establish for the first time a link between molecular biology and pattern formation. When this text is revised, the insect chapter at least will require complete reorganization. And that will be a measure of progress. □

Michael Akam is a Medical Research Council Senior Fellow in the Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK.

Green revolution

Donald Boulter

Principles of Plant Biotechnology: An Introduction to Genetic Engineering in Plants. By S.H. Mantell, J.A. Matthews and R.A. McKee. Blackwell Scientific: 1985. Pp. 269. Pbk £10.80, \$19.95.

PROGRESS in plant genetic engineering and in development of the potential applications has been faster than was anticipated. In these circumstances it is clearly difficult to write an account of the subject suitable for undergraduates because it is the latest advances which are of most interest. In the event, this book is good on the fundamentals of plant molecular biology but this aspect has been over-emphasized at the expense of the biotechnology. Moreover, the subtitle, *An Introduction to Genetic Engineering in Plants*, is confusing because many other areas of plant biotechnology are covered.

This is a somewhat bland account of an exciting field. The authors have been too equivocal in dealing with their subject matter, so that in parts they have been rather long-winded. The book is also too academic, no doubt reflecting the authors' experience, and often lacks critical assessment of the likely practical outcome or potential economic development of plant

biotechnology; for example, in the chapter on crop breeding, the section on the development of nitrogen-fixing cultivars in non-leguminous crops fails to address the whole area of the possible environmental impact of the release of genetically engineered organisms in relation to the principles and concepts of population biology and microbial ecology. Other topics inadequately dealt with include ice nucleation bacteria, herbicide resistant plants, and pest and disease resistance.

The book covers the whole field of plant biotechnology and contains basic information on genes, genetic engineering, tissue culture, cryopreservation, crop improvement by somaclonal variation and genetic engineering, plant-derived biomass and industrial plant products. The authors have read widely, and with the reservations made earlier most of the relevant information is included with relatively few points on which I would disagree. Unavoidably, however, some of the information is now out of date.

For all its failings, this book is a good attempt to cover a difficult topic. The readership at which it is aimed, undergraduates and scientists in related fields, will find it a helpful guide to what is going on in plant biotechnology, especially on the molecular biological side of things. □

Donald Boulter is Professor in the Department of Botany, University of Durham, Durham DH1 3LE, UK.

Meeting parasites

F.E.G. Cox

Foundations of Parasitology, 3rd Edn. By Gerald D. Schmidt and Larry S. Roberts. *Times Mirror/Mosby/Blackwell Scientific*: 1985. Pp.775. \$34.95, £33.50.

LIKE International Congresses of Parasitology, new editions of *Foundations of Parasitology* appear every four years. This is probably the most popular textbook on the subject, and is now bigger and better than ever before. It is a conventional and comprehensive account, in the tradition of Chandler and Read, and Noble and Noble, and deals in some detail with the parasites of man, in somewhat less detail with those of domesticated animals and in varying detail with those parasites of more interest to zoologists. Starting with definitions, principles and concepts, the 40 chapters work their way through the protozoa, helminths, pentastomids, crustaceans, insects and arachnids in a logical manner, leaving very few obvious gaps.

Within this largely systematic approach, the authors have tried to incorporate some of the modern aspects of immunology, biochemistry and physiology, and have done so very successfully by introducing just enough of these subjects to maintain interest without interrupting the logical sequence of the text; epidemiology and parasite ecology, however, receive scant treatment. The text is enlivened by copious drawings and photographs. These are mainly taken directly from the literature, which means that they are very variable in quality; still, it is pleasant to encounter so many old acquaintances. There are also ample references (though somewhat arbitrarily chosen), and students will welcome the fact that key words are printed in bold type and will find the glossary indispensable.

Progress in parasitology makes a book such as this impossible to keep up to date, but there can be no excuse for perpetuating the myth that in the malaria life-cycle merozoites reinfect liver cells as shown in Fig.8.1 (the text gets the story right, however). There are a few other places where newer figures would have been appropriate, and perhaps reference to some of the many recent reviews would have been better than some of the outdated papers cited. However, any criticisms are minor given the strengths of the book, and I hope that the authors will feel able to undertake the task of preparing a fourth edition in time for the 1990 International Congress. □

F.E.G. Cox is a Professor in the Department of Biophysics, Cell and Molecular Biology, King's College London (KQC), Strand, London WC2R 2LS, UK.

All about the mosses and liverworts

A.J.E. Smith

Introduction to Bryology. By W.B. Schofield. *Macmillan/Collier Macmillan*: 1985. Pp.431. \$45, £45.

DURING the past ten years several books and conference reports on the Bryophyta have been published, but this is the first textbook in the classical systematic botany genre to appear for some time. Almost two-thirds of it is devoted to detailed, excellently and abundantly illustrated descriptions of the morphology and anatomy of representative members of the various subclasses of mosses and orders of liverworts. Each account usually begins with an itemized list of distinguishing features, followed by descriptions (which contain a great deal more information than is readily available elsewhere) and finally a section on evolution and relationships. As is inevitable in writing about a group of plants about whose ancestry nothing is known, some of the remarks about evolutionary trends, either inferred or stated, are very much open to question. This is not pointed out in the text and what in many instances are mere hypotheses are likely to be taken as gospel by the unwary reader.

Most of the remainder of the book is given over to brief chapters on history, cytology and genetics, chemistry, physiology and ecology, topics which are covered, in most instances in more detail, in D.H.S. Richardson's *The Biology of*

Mosses (Blackwell Scientific, 1981). The last chapter, dealing with world distribution patterns, is longer and clearly the subject matter is closer to the author's heart.

Each chapter ends in a very adequate list of further reading, but there are no references in the text. Whilst this makes for a flowing style, it can be extremely irritating not knowing in which of up to 110 references additional information may be sought. This criticism applies particularly to the later chapters because of their introductory nature.

Also included are a glossary, and ten appendices covering topics ranging from a world list of floras, collecting and keys to major taxonomic groups, to squash techniques for the study of moss chromosomes (too inadequate to be of use). The contents of most of these appendices could have been incorporated into appropriate parts of the text with a considerable saving of space.

The work is more lavishly produced than is usual in specialized textbooks and there are few typographical errors. The style of writing is excellent and the author's enthusiasm is abundantly clear, especially in the systematic section which is at the same time easily intelligible to the beginner and readily acceptable to the knowledgeable. The book should appeal to the field biologist who wishes to know more about mosses and liverworts, as well as to advanced students of bryophyte systematics, and will serve as a reference book for those interested in other aspects of bryology. □

A.J.E. Smith is Reader in the School of Plant Biology, University College of North Wales, Bangor, Gwynedd LL57 2UW, UK.

On the waterfront

Ralph A. Lewin

The Physiological Ecology of Seaweeds. By Christopher S. Lobban, Paul J. Harrison and Mary Jo Duncan. *Cambridge University Press*: 1985. Pp.242. £30, \$44.50.

PROMETHEUS, who upset the Gods by stealing some fire from the Sun, was condemned to being chained to a cliff, where he had to endure the rigours of a harsh Caucasian climate. Many seaweeds, perhaps in expiation for their utilization of solar energy, have to endure an equally uncomfortable existence. Baked by unshaded suns in summer, scourged by icy waves and even chillier winds in winter, they are not to be envied. How do they manage to survive? This and similar problems are the subjects of *The Physiological Ecology of Seaweeds*. Lobban and his colleagues have done a wonderful job. The old saw about "filling a much-felt need"

has been misused often enough, but in this case its use is fully justified.

Physiological ecology is, I suppose, a branch — perhaps the main leader — of synecology. For this book, the title has been interpreted generously, and there is also a fair amount of biochemistry here, as is only proper. Light and photosynthesis occupy 30 pages; carbon metabolism, 20; nutrients, 35; temperature and salinity effects, 24; water motion, 14. All that takes up about half of the book. There are 34 pages dealing with communities of the seashore, where land meets sea and autecology meets synecology. (One may question the relevance of sections on grazing, species diversity and population dynamics, topics which are certainly ecological but hardly physiological.) Some 20 more review morphogenesis, the physiological interactions in multicellular systems. Finally, 20 pages are devoted to pollution: practical matters such as spilled oil and other organic and inorganic waste products of improvident mankind. Not much of relevance has been omitted —

PLENUM: TRACKING THE FRONTIERS

VIDEO MICROSCOPY

by Shinya Inoué

A practical handbook for those getting started in video microscopy. The basics of both video and microscopy are reviewed and sufficient theory is provided to aid in the selection and use of video microscopy equipment. A list of hard-to-find pamphlets—along with the addresses of their sources—and a glossary—with relevant section numbers included for each entry—are among the many features that make this a valuable reference as well as a practical handbook.

0-306-42120-8/528 pp. + index/ill./1986

\$65.00 (\$78.00 outside US & Canada)

text adoption price on orders of six or more copies: \$37.50

LEARNING ABOUT ENERGY

by David J. Rose

A wide-ranging text for senior and graduate students that weaves technical and scientific information into an integrated current account of the field, including energy in its social and economic contexts, environmental consequences of energy use, energy conservation, fossil fuel resources and reserves, coal, nuclear power, solar power, and global electrification. A volume in the series *Modern Perspectives in Energy*.

0-306-42124-0/484 pp. + index/ill./1986

\$59.50 (\$71.40 outside US & Canada)

text adoption price on orders of six or more copies: \$39.50

HYBRIDOMA TECHNOLOGY IN THE BIOSCIENCES AND MEDICINE

edited by Timothy A. Springer

"Altogether this is an interesting volume, containing much useful reference material about new ways to produce and improve monoclonal antibodies." —*Nature*

0-306-41996-3/628 pp./ill./1985

\$75.00 (\$90.00 outside US & Canada)

text adoption price on orders of six or more copies: \$45.00

MOLECULAR AND CELLULAR MECHANISMS OF ANESTHETICS

edited by Sheldon H. Roth and
Keith W. Müller

This comprehensive collection reports recent progress in this area, summarizes the current situation, and indicates future directions for research.

0-306-42128-3/483 pp. + index/ill./1986

\$59.50 (\$83.40 outside US & Canada)

Plenum Publishing Corporation

233 Spring Street
New York, N.Y. 10013

In the United Kingdom:
88/90 Middlesex Street
London E1 7EZ, England

Reader Service No.2



certainly there is nothing of importance that the authors haven't at least touched on. The book deals almost exclusively with seaweeds, considerations of marine angiosperms and of phytoplankton having been deliberately excluded.

Things are explicitly set out, with well-chosen diagrams, clear photographs and lots of charts and formulae occupying almost as much space as the text itself. A thousand references are quoted, in full, most of them from the past 20 years. The sentences are short and lucid; proof-reading has been meticulous; the index is adequate; the cover design is novel and attractive.

Not all is perfect, however. The legends of Fig. 10.8, a and b, seem to have been transposed. I think it would have been fairer, on p.137, to express concentration factors on the basis of wet weights (or even volumes), rather than on dry weights, be-

cause the seaweeds are compared, bulk for bulk, with volumes of water (which is, after all, wet). Figure 4.12 — illustrating changes in the monthly temperature and salinity coordinates for two locations — would have been easier to interpret if the months had been specifically indicated. And I wouldn't have called laminarin a starch, or methane a component of crude oil. These, however, are small matters.

The information and ideas in this book have never before been brought together in such a concise, readable form. The authors are to be congratulated for having accomplished a herculean task (a not inappropriate analogy, by the way, for it was Hercules who finally liberated Prometheus from his cliff). □

Ralph Lewin is a Professor of Marine Biology at the Scripps Institution of Oceanography, University of California, La Jolla, California 92093, USA.

Fringe existences

Stephen Hawkins

An Introduction to Coastal Ecology. By P.J.S. Boaden and R. Seed. *Blackie/Chapman & Hall:1985. Pp.218. Hbk £19.95, \$39.95; pbk £9.95, \$19.95.*

AN INTEGRATED treatment of coastal ecology at undergraduate level has been long overdue. In this book, Boaden and Seed have used their world-wide research and teaching experience to weave together the diverse strands of ecological research on coasts into a generally thorough yet approachable narrative.

The scope is admirably broad, moving from the edge of the continental shelf, via shores and estuaries, to dunes and coastal heath. An introduction to the coastal environment is followed by chapters on coastal habitats and their organisms: rocks, sediments, brackish waters, coral reefs, marshes and mangroves are all well covered, with "Other Coastal Habitats" completing this part of the text. The book finishes by dealing concisely and interestingly with inshore resources, seabirds and coastal management. A comprehensive bibliography lists key papers and reviews.

The book's balance, tilted towards marine ecosystems and animals, reflects the interests of the authors and there is, perhaps, not enough on the terrestrial systems consigned to "Other Coastal Habitats". Perhaps, too, the initial description of the coastal environment could have stressed its characteristic features more strongly. The sharp environmental gradients up and along shores and estuaries, and across dunes, mangroves and marshes, are insufficiently emphasized; for example, the time spent in or out of water at various tidal levels is relegated to a pa-

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Atlantic puffin (*Fratercula arctica*). The picture is taken from *The Atlantic Alcidae*, published by Academic price hbk £34, \$40; pbk £16.50, \$19.95.

graph under the subtitle "Seiches". The variability and unpredictability of coastal systems could also have been highlighted, and it would have been helpful if the different accounts of habitats and organisms had been drawn together in a synoptic overview. Conversely, the authors have done well to bring the study of seabirds into the marine biological mainstream, while the chapter on the neritic province provides a description of basic biological oceanography which beginners will find particularly illuminating.

This is a useful textbook which is easy to read and understand. It should be helpful to non-biologists interested in coastal processes, and will be appreciated by undergraduates embarking on options or field courses in marine ecology; certainly, it provides a firm basis for subsequent, more advanced reading. My opinion of the book is best summed up by the fact that I have adopted it for my own course on coastal marine biology. □

Stephen Hawkins is a Lecturer in the Department of Zoology, University of Manchester, Manchester M13 9PL, UK.

The why of what was and is where

Carolyn Harrison

Basic Biogeography, 2nd Edn. By Nigel Pears. Longman:1985. Pp.358. Pbk £9.50, \$17.95.

WHEN *Basic Biogeography* first appeared in 1977 it made its mark as a text that was readable, lucid and sensitive to the demands of the learning process. In this second edition the same structure is retained and revision has involved modest addition rather than comprehensive recasting.

Part 1 introduces basic concepts and terms as they apply to vegetation, soils and ecosystems. The approach is essentially an ecological one, rather than chorological or evolutionary, and through it the student will gain a rudimentary knowledge of some of the complexities exhibited by natural and disturbed ecosystems. Part 2 stands largely on its own and provides an account of selected examples of the vegetation and soils of Britain. A final

chapter describes the impact of man in terms of land-use history and land-use conflict.

The addition of topics such as homeostasis, plant strategies, island biogeography and carrying capacity, as well as a teaching model of nutrient cycling, strengthens the book as an introductory text, while the new, bolder type lends added crispness to the page. Appealing, too, are the discussion sections at the end of each chapter; these serve to encourage a more critical approach to the subject material, which otherwise is presented in a descriptive fashion. Even so, it is to be hoped that inquisitive students will be moved to examine more fully the ecological basis of the concepts and theories upon which the book is founded. If they do not, many of them will remain ignorant of the true extent of controversy which each topic generates amongst ecologists. On balance, however, I am optimistic that the new edition of *Basic Biogeography* will provide just the right kind of spur to dig deeper into the subject that most students need.

Carolyn Harrison is a Lecturer in the Department of Geography, University College London, 26 Bedford Way, London WC1E 0AP, UK

Ecology writ large

Robert H. Smith

Ecology: The Experimental Analysis of Distribution and Abundance, 3rd Edn. By Charles J. Krebs. Harper & Row: 1985. Pp. 800. Hbk \$32.50; pbk £12.95.

Ecology: Individuals, Populations and Communities. By Michael E. Begon, John L. Harper and Colin R. Townsend. Blackwell Scientific/Sinauer: 1986. Pp.880. Hbk £29.80, \$33.95; pbk £14.50. To be published in April.

CAN A single textbook now give an adequate account of the science of ecology at university level? In the 1950s and 1960s, comprehensive coverage of the subject was achieved in the successive editions of Odum's *Fundamentals of Ecology* (first published by Saunders in 1953), which served as the standard text for many years. More recently, however, there has been a trend among publishers to produce series of shorter books written for students by a specialist in a particular field, perhaps indicating the extent to which ecological theories and ecological studies have blossomed and borne fruit during the past two decades.

Two substantial textbooks dominated undergraduate reading lists during the 1970s: Ricklefs' *Ecology* (Nelson, 1973) and Krebs's *Ecology* (Harper & Row, 1972). Krebs's book has now run to the third edition and must still be considered

the leader among general ecology texts. It is against this competition that the new-comer from Begon, Harper and Townsend must be judged. Both books are some 800 pages in length (compared with about 500 pages of text in the 1959 edition of Odum), and include an extensive, up-to-date list of references to original papers. Both, too, justify the answer "yes" to the question posed at the beginning of this review.

The flavour of Krebs is brought out in the full title, *Ecology: The Experimental Analysis of Distribution and Abundance*. The emphasis on the experimental approach has raised ecology above the level of observational natural history, and is one of the main features of Krebs's own research on small mammals. Hypotheses and comparisons are presented throughout the book, and readers should be left in no doubt that ecology is an area of science which, though not in the same "hard science" league as physics and chemistry, has firmed up considerably in recent years.

Krebs has taken account of the main developments in ecology since the earlier editions and has produced a substantially revised volume. Gone are Chapters 26 and 27, "Human Population" and "Food Production"; brought forward and given more prominence is "Evolution and Ecology" which sets the scene for new sections such as those in the chapter on habitat selection; much increased is the coverage of community ecology and species diversity. Of course, a reader wanting to find out

about a particular topic cannot expect to find everything in this or any other textbook. Many sections are tantalizingly brief, but throughout Krebs refers to the primary literature and there is a selected reference list in each chapter as well as a full bibliography at the very end. Overall, those who already know Krebs's *Ecology* will find the new edition a mature version of an old favourite, while those new to the subject still have something of a treat in store.

Begon, Harper and Townsend sub-title their book *Individuals, Populations and Communities*, which describes the logical progression that they follow through the 22 chapters. Their approach integrates almost all areas of ecology, while recognizing that there are different (but interconnected) levels at which the subject can be studied. Appropriate examples are discussed throughout, and the authors mix plant and animal examples in a wholly natural way; the distinction they make (stemming largely from Harper's seminal work of the past two decades) is between unitary and modular organisms rather than animals and plants, a classification which would be inappropriate in most areas of biology but is probably the best one for ecology. The argument that the characteristics of individuals are the products of natural selection, and that the characteristics of populations depend on the individuals that make up those populations, is emphasized in a number of places, and rightly so; the authors thereby avoid leading students towards the pitfall of implicit group selection.

Indeed, the book is commendable in almost every respect, though there are some points to quibble over. I found the three introductory chapters rather ponderous, though much of the material is important. I have reservations about the emphasis on reproductive value in the chapter on life history variation, because the authors do not make clear how the relative fitnesses of alternative life histories should be compared using reproductive value. Also, the explanation of the use of zero isoclines to analyse stability of predation dynamics is fudged somewhat — crystal clear for competition but not so for oscillating populations, and students may well be confused.

As a whole, though, the book is excellently written, well-presented and emphasizes the areas of ecology currently attracting most research interest. The freshness of approach and the breadth and depth of coverage make most of the opposition seem a rather stale and restricted diet by comparison (Krebs excepted). All ecology teachers and students would do well to sample this *nouvelle cuisine*. □

Robert H. Smith is a Senior Lecturer in the Department of Pure and Applied Zoology, University of Reading, Whiteknights, Reading RG6 2AJ, UK.

Populated classes

Michael B. Usher

Lectures in Theoretical Population Biology. By Lev R. Ginzburg and Edward M. Golenberg. Prentice-Hall:1985. Pp.246. Pbk \$19.95, £21.35.

Case Studies in Population Biology. Edited by L.M. Cook. Manchester University Press/Longwood Press, 51 Washington St, Dover, NH 03820:1985. Pp.218. Hbk £18.95, \$35; pbk £7.50.

THE expression "population biology" is now applied in a variety of ecological and genetical investigations. The two books being reviewed, although containing the same two words in five-word titles, are totally unlike each other, and could be considered to be in opposing corners of the plane that represents the subject. One views population biology largely mathematically, and introduces it via genetics; the other views it almost as precise natural history, and introduces it via the results of undergraduate field classes.

Lectures in Theoretical Population Biology is arranged in the form of 22 lectures, with a short problem-solving section and homework exercises after every three or four lectures. The authors have taught this course at the State University of New York at Stony Brook, and they say "It was, for all the students, their first exposure to population biology covering areas of evolution and ecology, and for the majority of them it was also their last". A certain amount of knowledge would be required to start a course such as this, though explanations, especially in the first few lectures, are good. The approach is initially genetical, with consideration of Hardy-Weinberg laws and of selection in natural populations leading to topics such as mutation, inbreeding and drift in small populations, for all of which the opportunity has been missed to stress the applications in such areas as wildlife conservation. The second half of the book is ecological, with investigations of population increase, predation, mathematical models and their application in demography. There is an attempt to bring the two halves of the book together in the concluding lecture that views "fitness" as the common currency.

For two decades students from Manchester, and to a lesser extent from Liverpool, have visited Woodchester Park, a valley in south-west England, on summer-time field courses, and a few graduates have undertaken their research projects in the area. Five of the six chapters in *Case Studies in Population Biology* report a variety of thorough research studies against a background of long-term, less-intensive work undertaken by generation after generation of undergraduates at

Woodchester. The remaining (opening) chapter describes the development of the practical field-testing philosophy and the locality in which the studies were undertaken, and provides summary results for some of the long-term monitoring (for example of damselflies, grasshoppers, moths and earthworms) and the relationships between the species' abundance and diversity on the one hand, and various environmental influences on the other. Subjects of the detailed studies are mice, birds and the scarlet tiger moth, as well as some of the groups mentioned above.

The approaches in the two books could not be more different. *Lectures in Theoretical Population Biology* is mathematically based and such data as are used are generally artificial (or, at least, hardly appear to be real). The subject is taught theoretically, and those students that remain interested (the quotation above

implies a large drop-out rate) can apply their knowledge in the real world. *Case Studies in Population Biology* starts in the real world by introducing students to ideas of diversity and abundance in many groups of animals. A few students continue to the PhD level, and the later chapters of the book indicate how theoretical concepts are used to solve real-world problems. While the Manchester book contains many references to follow up, the Stony Brook book contains none.

Either volume could help in the planning of a course in population biology, once the approach to the subject has been decided. However, for one of them a teacher will need a decade or two's data and for the other an element of mathematical competence in the class. □

Michael B. Usher is Reader in the Department of Biology, University of York, York YO1 5DD, UK.

Past imperfect

Beverly Halstead

Palaeontology: An Introduction. By E.W. Nield and V.C.T. Tucker. Pergamon: 1985. Pp.178. Hbk £15.95, \$20.75; pbk £7.95, \$10.35.

FOR SOME curious reason it seems impossible for anyone to put together a textbook on palaeontology for beginning students. I contemplated such a task at one stage but had the sense to give up before it was too late.

The authors of this book clearly have more courage and theirs is indeed a valiant attempt. The style is exceptionally easy to read. There are a number of telling analogies that will stay in the memory, such as the example of recent hatchback cars from different motor companies that end up the same "in consideration of aerodynamic function in body design"; perfect examples of homeomorphy. And how nice it is, too, to see a section entitled "Welcome to Palaeontology". So one cannot but feel something of a cad at not being able to go overboard with enthusiasm at their efforts.

One of the drawbacks of the book is the astonishing variety of inclusions and exclusions. There is a succinct discussion on "What is 'numerical taxonomy'?" and an illustration of a "dendrogram". Especially at a time when every child that visits the Natural History Museum in London, or sees their publications on fossil man or dinosaurs, will have had cladistics presented in no uncertain manner, one wonders where the authors were when cladistics was a major talking point in the media. Then there are the odd snippets such as ammonite aptychi, paired calcite shells, badly described as operculae with no hint

of the recent papers claiming them to be jaws. Punctuated equilibrium and phyletic gradualism rate a brief mention, but what of the new battle cry of "punctuated gradualism", which has emerged over the past few years? Among the vertebrates, the acanthodians now generally thought to be the stem of the bony fishes are nowhere to be found. The sense of being unnecessarily out of date pervades the book.

It is, too, a real puzzle to try and understand the logic of the sequence of subjects. The authors start with trilobites followed by graptolites, brachiopods, molluscs, echinoderms, corals and stomatopods, vertebrates, microfossils and finally vascular plants.

One of the excellent ideas is the series of diagrams describing the basic morphology of the main groups, each of which is followed by illustrations of common fossils that the student is likely to meet and to be asked to draw. Here the figures are rather variable. There is really a lot of information in them but although the notion is excellent their execution leaves a lot to be desired. The drawing of the ostracoderm *Psammolepis* suggests that the artist did not know where the mouth was situated; the figure of the poor lamprey is a travesty.

The authors have a flair and considerable ability to communicate their subject, but somewhere along the line they have been let down. If they brought their material up to date, put the chapters into some kind of logical sequence and got the scale and type of illustration right, then they could produce a truly excellent introduction. I only hope they will be encouraged to do so, for that would be a fine and useful book. □

Beverly Halstead is Reader in the Departments of Geology and Zoology, University of Reading, Whiteknights, Reading RG6 2AB, UK.

Putting the backbone into zoology

Tom Kemp

Vertebrate Life, 2nd Edn. By William N. McFarland, F. Harvey Pough, Tom J. Cade and John B. Heiser. *Macmillan/Collier Macmillan*: 1985. Pp.636. Hbk \$38.95, £35; pbk \$18.90, £14.50.

Looking at Vertebrates: A Practical Guide to Vertebrate Adaptations. By Elizabeth Rogers. *Longman*: 1986. Pp.195. £9.95.

THE first edition of McFarland *et al.*'s *Vertebrate Life* appeared in 1979, and was a most refreshing textbook. It consisted of a comprehensive review of the vertebrates as living organisms, dealing with their major biological features as adaptations for their respective ways of life, and integrating into the text accounts of their fossil history and the diversity within the main groups. While this approach was exemplary and largely successful, there were weaknesses, principally in the rather dated, unsophisticated discussion of the fossils and their relationships to the living members.

In the new edition this deficiency has been largely remedied. There are still a number of outdated ideas presented, such as a phylogeny of early mammals which shows an extensively polyphyletic origin (and contradicts another phylogeny, of synapsid reptiles, showing a monophyletic origin of mammals). Nevertheless, such criticisms count for little against the positive achievements of the book.

The organization is traditionally phylogenetic, running from basic plan and origin, through agnathan fishes, jawed fishes, amphibians, reptiles, birds and mammals. The discussions of the fossil groups are confined to separate chapters, but there is adequate cross-referencing to the chapters about the living forms. There are also separate chapters on the geology and ecology of the Eras, in which are to be found modern, competent accounts of such things as palaeoclimates and continental drift. Many other small, but interesting topics are mentioned throughout the book, such as the amount of cellular DNA in different groups, and current ideas and controversies about evolutionary mechanisms get an airing. These passages are stimulating, even though the subject matter is not always developed enough for the relevance to be entirely clear.

Naturally, many topics are dealt with sketchily, and the student will need to extend his reading. It is therefore a great help that the reference lists have been made more comprehensive and up to date, particularly as far as review-type papers are concerned. All in all, I shall

continue to regard this as the best basic textbook to accompany my own lectures on vertebrate zoology.

Looking at Vertebrates is a rather unusual practical book. Each of the eight chapters is, in essence, a practical class in which certain aspects of a particular group of vertebrates are studied. For example, a comparison of elasmobranchs and teleosts, a study of amphibian locomotion, and a survey of mammal feeding constitute parts of three respective chapters. In each case, living animals are discussed, directions and illustrations for the appropriate dissections offered, and the morphology of the relevant skeletal elements, muscles and so on noted. The emphasis throughout is on the adaptive significance of structure, and to this end

the text is littered with the "Why is it like that?" kind of question that demonstrators traditionally ask of their students.

Because there is no attempt to be comprehensive, this is, by its nature, an idiosyncratic book. Experienced lecturers will generally prefer to select their own topics (although they may well gain inspiration from the general approach adopted here), but for the new teacher who is landed with having to organize a series of vertebrate practicals here is a ready-made course. Certainly the book is nicely illustrated and the enthusiasm of the author for her subject is unmistakable.

Tom Kemp is Curator of the University Museum, Parks Road, Oxford OX1 3PW, and a Lecturer in the Department of Zoology, University of Oxford.

Course of change

A. Hallam

Earth and Life Through Time. By Steven M. Stanley. *W.H. Freeman*: 1985. Pp.690. \$35.95, £39.95.

WITHIN the past few years, developments in the earth sciences have made possible a satisfactorily integrated view of plate tectonics and climatic and oceanographic changes in relation to biogeographic events and organic evolution and extinction. Stanley's ambitious book is intended as a self-contained introduction to both biological and geological aspects of our planet's evolution. Unlike many other textbooks dealing with historical geology, it demands no previous knowledge of physical geology.

After an introductory chapter, the early part of the book describes the environments that can be inferred from the study of modern and ancient sediments and their organic contents, with correlation and rock dating, and with the history of life inferred from the fossil record. Subsequently there are chapters on plate tectonics and mountain building. The concluding, most substantial, part of the book is devoted to historical geology, the evolution of the Earth from the Archaean to the present. Several useful appendices on rocks, minerals, fossils and stratigraphic stages are included, and there is also a glossary.

In its assumption of little prior knowledge, the abundance of photographs and diagrams, and inclusion of chapter summaries, exercises and modest reference lists, *Earth and Life Through Time* is clearly aimed primarily at the large North American elementary undergraduate market. The bias towards North American geological examples is not so great, however, as to make the book of limited value to students in other continents.

In a work so wide-ranging in scope it would be remarkable if there were no variability in quality. As might be expected, Stanley is excellent and thoroughly up to date on organic evolution, including the early development of life in the Precambrian, and in this respect the book is far better than its rivals. Rather more perfunctory in his treatment of sedimentary environments, and his account of tectonics and mountain building pays almost no attention to the excitement stemming from the concept of collage tectonics. Thus many of the Pacific borderlands and large areas of eastern Asia are now thought to have been progressively accreted since the Palaeozoic as a series of "suspect" or "displaced" terranes. It is regrettable that so much reference is made to the familiar Exxon eustatic curve published in 1977, because the "sawtooth" shape has been disowned by Vail and his colleagues for several years as reflecting coastal onlap patterns rather than sea level. In his account of the possible causes of mass extinctions Stanley shows an undue bias towards temperature control, which many of us think of as being only one of a number of causal factors, and by no means the most important during the long stretches of time when the Earth apparently enjoyed an equable climate.

For a book so clearly written and well illustrated, however, and which gives a sound introduction to so many topics, it would be churlish to quibble over such points. Rather, one should express the hope that it will be widely adopted as a course text. □

A. Hallam is Lapworth Professor of Geology, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

● *Time Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria*, by Niles Eldredge, will be published in Britain next week by William Heinemann. Price is £12.95. The book was reviewed by Richard Dawkins in *Nature* 316, 683 (1985).

Oil: where it is and how to find it

R.C.L. Wilson

Petroleum Geology. By F.K. North. *Allen & Unwin*:1985. Pp.607. Hbk £50, \$60; pbk £19.95, \$34.95.

Elements of Petroleum Geology. By Richard C. Selley. *W.H. Freeman*:1985. Pp.449. \$44.95, £19.95.

Geology in Petroleum Production: A Primer in Petroleum Geology. By A.J. Dikkers. *Elsevier*:1985. Pp.239. Dfl.120, \$44.50.

PETROLEUM geology is a vast field to cover in a single textbook. Petroleum geologists must have a good grounding in many aspects of geology: palaeontology, stratigraphy, organic geochemistry, sedimentology, structural geology and geophysics. They must also be familiar with the principles — though not the practice — of engineering related to the hydrocarbon exploration industry, as well as relevant legal and economic matters. Each of these three books sticks firmly to the geological side of the subject.

North and Selley invite immediate comparison. Both are aimed at advanced students and both cover the same topics, albeit in different orders and under different headings.

Part I of North is a broad introduction, discussing, in four short chapters, what petroleum geology is, its basic vocabulary (a two-page chapter which seems pointless), some basic statistics (from units used in the industry, through people to wells, fields, basins, reserves and costs), and, finally, the history of petroleum exploration from 1942 onwards. Selley's introduction is much shorter and to the point. He entitles his historical review "Petroleum from Noah to Opec", and here also lists the conditions necessary for a commercial oil accumulation to occur. A clear exposition of these is not given by North until much later on in his book, and altogether his "signposting" in the introductory part — essential if a student is to be encouraged to read on — is less effective than Selley's.

The rest of *Petroleum Geology* appears to be structured in a logical way, commencing with the nature and origin of petroleum (Part II), then moving on to how it accumulates (Part III), followed by eight chapters on exploration and exploitation (Part IV) and concluding with a summary of the distribution of oil and gas (Part V), which includes case histories of selected fields. However Part III (where and how oil and gas accumulate) is rather an eclectic mixture, dealing in various chapters with porosity and permeability, well logs, reservoir rocks and pressure conditions

within them, migration of oil and gas, and trapping mechanisms.

The ordering of subject matter in *Elements of Petroleum Geology* is rather more satisfactory, for having read about the physical and chemical properties of oil and gas in Chapter 2, the reader knows what he is looking for, and can then learn how to find it in the following chapter on methods of exploration. Here Selley reviews drilling techniques, wireline logs, geophysical exploration methods, subsurface mapping and remote sensing, before moving on to examine the subsurface environment under the headings of waters, temperatures, pressures and fluid dynamics. The reader is thus equipped to consider how such environments may lead to hydrocarbon generation and migration, and to get to grips with the nature of reservoirs and of traps, in the following chapters.

Both authors deal at length with the form, distribution, origins and classifications of sedimentary basins. Although both of them relate sedimentary basin formation to plate tectonics, neither explicitly claims that the new geology has helped find oil. Indeed, North states (p.351) that "The advent of plate tectonics did surprisingly little to alter the geological understanding of basins. It provided a unified framework within which to identify the appropriate places, and it made a chaotic basinal nomenclature even more so". The last statement seems to be going a bit too far — especially as Selley shows that a nice tidy picture can be presented within a plate-tectonic framework.

The only major omission from these

two books is a treatment of the aspects of structural geology especially relevant to petroleum exploration geology. Recent advances in the study of thrust belt terrains and areas of extensional tectonics are not discussed. This is a pity, because we are now beginning to see, through deep crustal reflection seismology, how structures formed during the compressional demise of one basin may influence the development of extensional features that create successor basins.

More surprising is the complete absence of reference lists in North, save for a brief review of the literature in the first chapter. The only references given are authors and journal name and year of publication in sources quoted in the figure captions. This failing is incomprehensible, and is made worse by the frequent name-dropping in the text. In contrast, each chapter of Selley concludes with a selected bibliography and a full list of references cited.

In all, *Elements of Petroleum Geology* is an excellent, very readable text for final-year undergraduates. North's book contains much more detail, as reflected in its greater length, and so would be more suitable for supporting work on a master's degree and for graduates newly embarked on a career in petroleum exploration.

The market for Dikkers's *Geology in Petroleum Production* is less easy to define, though petroleum production is clearly a postgraduate or vocational subject. In his preface, the author states that his book has a "comic strip" style; to a degree it has, for each topic is treated via a few pages of text and a few unattractive diagrams.

The book is divided into 11 parts (including an introduction and epilogue) which deal with correlation, tectonics, reservoir geology, reservoir fluids, accumulation conditions, appraisal drilling, development drilling, reserve estimates and a field review study. Unfortunately, most of the examples used are fictitious, rather than of actual "geological conditions found in nature". The author's justification for this is that "apart from the administrative and legal difficulties involved in collecting such evidence, it is rare that Mother Nature presents us with examples which illustrate with satisfactory clarity the phenomenon one wants to demonstrate". This is hard to believe, especially after seeing the "real" examples to be found in Selley and North.

Given this approach, and the fact that most of the topics are covered both by the other two books reviewed here and by free manuals issued by exploration service companies, I am puzzled about who might wish to purchase a book which is less than half the size of *Petroleum Geology* but costs much the same. □

R.C.L. Wilson is a Senior Lecturer in Earth Sciences at the Open University, Milton Keynes MK76AA, UK.



Eruption of Cerro Negro, near Managua, Nicaragua, in 1968. The picture is taken from the new fourth edition of *Earth*, by Frank Press and Raymond Siever, published by W.H. Freeman. Price is hbk \$32.95, £32.95; pbk £16.95.

Mark Hurd Aerial Surveys

On the rocks

K.G. Cox

Magmas and Magmatic Rocks: An Introduction to Igneous Petrology. By Eric A.K. Middlemost. Longman:1985. Pp.266. Pbk £14.95.

Igneous Petrology. By Alexander R. McBirney. Freeman, Cooper/Oxford University Press:1985. Pp.504. \$38.75, £37.50.

The Field Description of Igneous Rocks. By Richard Thorpe and Geoff Brown. Open University Press/Halsted Press:1985. Pp.154. Pbk £6.95; hbk \$13.95.

STUDENTS of igneous petrology are particularly well served by textbooks at present. Whether the explosion in publications at this level has been a consequence of the generally healthy state of the earth sciences, and specifically of igneous petrology, as a result of the rise of the X-ray fluorescence spectrometer and the electron microprobe, or whether it is due to erosion of academic salaries, is not clear. The fact remains that there are now at least half-a-dozen reasonably up-to-date and general texts to choose from.

Magmas and Magmatic Rocks by Eric Middlemost will find a place among the best of them. To give an idea of its particular flavour it is worth back-tracking to that classic of petrological writing *Igneous and Metamorphic Petrology* by Turner and Verhoogen, published in 1951 with a second edition in 1960. Turner and Verhoogen was strong on taxonomy, on grouping the diverse world of natural rocks into a small number of associations, and in digesting and presenting petrogenetic ideas. For those of us used to this view of the world, Middlemost is a familiar but very welcome newcomer. It is a rather short book, essentially in the Turner and Verhoogen style, but contains the information one would want to find. It is strong on descriptive details (particularly the geochemical) and presents sensible consensus petrogenetic ideas. It is less authoritative on magmatic processes, but so was Turner and Verhoogen, and it deserves to become a widely used and successful textbook.

Alexander McBirney's *Igneous Petrology* is an equally impressive but very different work. McBirney has always been in the forefront of the new wave of igneous geologists who are rather more concerned with how magmas behave than with what rocks are, an approach originating more in volcanology than in geochemistry and petrology. Times change, and the ideal student of the future will absorb what both Middlemost and McBirney have to say. However, McBirney is currently likely to be less popular. The book is long and rather expensive, with a high intellectual

content that places considerable demands upon the reader. Much of it is permeated by the concepts of thermodynamics and fluid dynamics, with many quantitative examples, culminating in an appendix of intriguing problems (answers are not given).

The committed student will find this book a gem, and it points the way that igneous petrology will undoubtedly take in the future. It is perhaps a pity, however, that the author has attempted to make the book comprehensive by including large amounts of descriptive information in a manner rather similar to Middlemost. These sections are written with less conviction than the others, and I would have preferred to see a shorter book with more concentration on McBirney's special talents.

The Field Description of Igneous Rocks is the latest addition to the Geological Society of London's *Handbook* series. It is a small, soft-covered yet tough, volume which all undergraduates embarking on their finals mapping project should carry in their knapsacks. Apart from containing the obvious things such as hints on how to keep a field notebook and on other specific field skills, it includes a great deal of general information on the topics of volcanology and the structures of plutonic rocks. It is thus a highly useful introductory text for those studying any aspect of igneous geology, and should have many applications in class work as well as in the field. □

K.G. Cox is a Lecturer in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

Settling standards

Philip Allen

Principles of Physical Sedimentology. By J.R.L. Allen. Allen & Unwin:1985. Pp.272. Hbk £30, \$40; pbk £14.95, \$24.95.
Experiments in Physical Sedimentology. By J.R.L. Allen. Allen & Unwin:1985. Pp.63. Pbk £7.95, \$9.95.

JOHN Allen's textbook, *Physical Processes of Sedimentation* (Allen & Unwin, 1970), was an early attempt to explain sedimentary mechanisms in terms of physical principles. In its insistence on a rigorous physical approach it echoed the tones of Bagnold's classic *The Physics of Blown Sand and Desert Dunes* (Chapman & Hall, 1954), and very much set the standard for students to follow. However, now over 15 years old, it was in need of revision and the present volume admirably succeeds in this task. The new book is accompanied by a helpful practical manual which describes 28 experiments illustrating some of the main principles of physical sedimentology.

In reality, little has occurred in the past 15 years with truly fundamental implications for the fields of loose boundary hydraulics and bedforms, with the notable exception of the development of ideas, germinated in the 1960s, on the nature of turbulence. The main differences between *Principles of Physical Sedimentology* and its predecessor is the inclusion of a new chapter on results of flow-visualization studies of turbulence and, the new vogue, the processes and mechanisms affecting muds.

Many students in sedimentology are daunted by equations describing processes which they do not have the basic knowledge to understand qualitatively, let alone quantitatively. The first chapter sets out to provide some of the basic rules of

fluid mechanics that can be used in the following 12 chapters, but it could easily have been twice as long, for it is these basics which provide the essential insights to sedimentology problems. For this reason I would have preferred a section at this stage on the structure of turbulent boundary layers rather than discovering it relegated to Chapter 6. Two further chapters, one considering the description of detrital particles and their packing, the other dealing with the settling of particles in a stagnant fluid (in effect an analysis of drag), complete this section on fundamentals.

Entrainment, sediment transport and the resultant bedforms in a cohesionless substrate should form the core of any book on physical sedimentology. Allen discusses this topic for unidirectional flows surprisingly briefly; equivalent information about oscillatory flows is found in the very last chapter and one has to wait tantalizingly for two chapters before a statement on suspended sediment transport is encountered. Even so, these individual chapters provide a marvellous physical background for the student.

Looming large in the second half of the book is a discussion of mass movements in general, with a concentration on gravity currents in the sea (turbidity currents). Additional chapters deal with the properties, deposition and erosion of muddy sediments and there is an interesting discourse on deformation of unconsolidated sediments.

Despite the curiosities in its organization and the idiosyncratic chapter headings, *Principles of Physical Sedimentology* will be an important text for students interested in the explanatory nature of sedimentology. It represents another fine contribution from a remarkable sedimentologist. □

Philip Allen is a Lecturer in the Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK.

Making of singular landforms

Paul W. Williams

Karst Geomorphology, 2nd Edn. By J.N. Jennings. Basil Blackwell: 1985. Pp.293. Hbk £27.50, \$49.95; pbk £9.50, \$14.95.

Limestone Geomorphology. By Stephen Trudgill. Longman: 1985. Pp.196. Pbk £12.95, \$19.95.

SECOND editions of successful books can be very disappointing. Not so *Karst Geomorphology*. The new edition reflects Joe Jennings's indefatigable energy, enthusiasm and scholarship. It retains the best features of the first edition, yet shows a thorough revision, and there can be little doubt that the book will retain its place as the pre-eminent introductory text on the subject.

Karst geomorphology is the study of landform development on carbonate and evaporite rocks. Unusual landforms and hydrology develop because of the high solubility of these lithologies in natural waters. Carbonate rocks alone cover about 10 per cent of the continents, and are particularly highly represented in some areas: they outcrop over 45 per cent of France, 33 per cent of Yugoslavia and 15 per cent of the United States. There are also more than two million square kilometres of karst in China, including an incredible tower karst, which, for Jennings, provokes "images of giants in drunken reel transfixed into immobility". It is one of the world's great landscapes and has been the subject of traditional Chinese art for centuries.

Jennings helps his readers to understand karst by following the same successful formula as in his first edition. Again, the book is a model of clarity, and the newcomer to the field will appreciate from the chapter headings that karst is a microcosm of geomorphology as a whole. After describing the rocks it develops on and the processes that mould it, Jennings discusses and explains the special features that give karst its singularity, among them sinking rivers, springs, caves, gorges, karren, towers and a range of enclosed depressions (dolines, uvalas, poljes and cockpits). Of the concluding chapters, one is entirely new ("Coast and Karst") and one from the first edition has been omitted ("Present State of Karst Geomorphology and its Value") — sensible decisions, be-

• *Themes in Geomorphology*, edited by Alastair Pitty and published last year by Croom Helm/Barnes & Noble, is a collection of reviews on the broad subject of geomorphology and includes a chapter on karst forms and processes. The book is intended to offer "a geomorphological aperitif for students of geography . . .". Price is £25, \$28.50.

cause the lack of the former and the content of the latter were both weaknesses of the original version.

Most of the book has been rewritten, the original Chapters 4 and 5 and 9 and 10 have been reversed in order, and more than 50 per cent of the extensive reference list is new, with entries to 1984. To those of us who had badgered Joe Jennings to produce a second edition, it is a great relief that he completed the work before his untimely death. His book will live on to educate another generation of karst geomorphologists; Peter Bull and Andrew Goudie deserve our thanks for seeing it through the press.

Limestone Geomorphology by Stephen Trudgill is the eighth in a series of geomorphology texts under the general editorship of Professor Keith Clayton. Its publication in the same year as Jennings's second edition inevitably invites comparison between the two books. Trudgill's preface accurately describes his purpose: "The book is . . . primarily a selective assessment of recent work on limestone erosion processes, indicating, where possible, how process and form may be related".

This at once suggests that the title is inappropriate, but it by no means detracts from the undoubted value of the contents.

Almost half of the text is devoted to geomorphic processes. Complex subjects are handled skilfully with a masterly blend of theory and field data. Chemical and biological processes in the soil, on rock outcrops and along the coast are particularly well dealt with, reflecting Trudgill's original research contributions. Much less successful and substantial is his treatment of morphology and physical processes.

I regard Jennings's book as balanced, comprehensive and authoritative, and Trudgill's as too uneven to stand alone as an introductory text. Although it is narrower in scope and more limited in reference material, however, *Limestone Geomorphology* is much superior on processes, and generally better illustrated and easier to read. These, then, are a valuable and complementary pair of texts, but I shall require my students to purchase only one of them. □

Paul W. Williams is Professor in the Department of Geography, University of Auckland, Private Bag, Auckland, New Zealand.

Seismological event

Peter D. Marshall

An Introduction to the Theory of Seismology, 4th Edn. By K.E. Bullen and Bruce A. Bolt. Cambridge University Press: 1985. Pp.499. Hbk £35, \$69.50; pbk £15, \$24.95.

UNTIL his death in 1976, Professor Bullen was gathering material for a new edition of his acclaimed textbook, *An Introduction to the Theory of Seismology*, three editions of which had appeared between 1947 and 1963. Since the early 1960s much progress has been made in seismology, due largely to a massive injection of funds for research aimed at improving our understanding of the physics of seismic sources and elastic wave propagation phenomena, with the eventual object of resolving problems concerned with earthquake prediction and the monitoring of nuclear test-ban treaties. Many of the advances have been on the theory side of the subject, particularly in the development of improved statistical and mathematical techniques such as inversion theory, making radical revision of the book a necessity.

It is appropriate that Bruce Bolt, Professor of Seismology at the University of California, Berkeley, and a former colleague and student of Bullen, should have accepted the task of modifying and updating the original text. Bolt has tackled the job with characteristic enthusiasm. The first ten chapters on topics such as elastic-

ity theory, vibrations and waves are retained with only minor changes, and some new material has been added on seismometry, but the rest of the book has been drastically restructured and, with all the new material, greatly improved. There are, for example, new chapters on the theory of earthquake sources, seismic wave propagation through the Earth and its anomalous regions, free oscillations, strong motion seismology, earthquake prediction and risk. Each chapter has been carefully prepared to provide a good grounding in the fundamentals of the topic concerned, and there are exercises for the reader to test his comprehension.

With the additional material and the enlarged reference list, the fourth edition is longer than its immediate forerunner even though some topics such as seismic prospecting and model seismology have, quite rightly, been dropped. In earlier editions the illustrations were not exactly prolific, but Bolt has rectified this shortcoming by providing numerous real seismograms and clear informative diagrams which have been carefully selected to complement the text.

Bolt has drawn on his wide knowledge of seismology and expertise as a university teacher to produce a valuable and well-presented textbook. It will at once fill the course needs of students and teachers of seismology and related subjects, and be very useful to professionals engaged in seismological research. □

Peter D. Marshall is a Principal Scientific Officer in the Seismology Division of the Ministry of Defence (Procurement Executive), Blacknest, Brimpton, Reading RG7 4RS, UK.

Getting into Earth

Marcia McNutt

Geophysics: The Earth's Interior. By Jean-Claude De Bremaecker. Wiley: 1985. Pp.342. \$29.95, £30.45.

FIRST the good news. De Bremaecker's introductory account of geophysics is pitched at a simpler level of mathematics than Stacey's *Physics of the Earth*, while still allowing a more quantitative grasp of the subject than is possible using Bott's *The Interior of the Earth*. The book is intended for undergraduate earth scientists with one year of calculus and physics, and the mathematical derivations are easy to follow. Interesting historical notes make the text quite readable. I applaud De Bremaecker's use of SI units, but wish he had remained committed to them — mgals, kbars and kg/cm² all crop up.

Each chapter is divided along the classic branches of geophysics and presents the physics of the problem, methods of solution, a brief description of instrumentation, plus the most recent results. This last feature is a mixed blessing. For a year or two at least it will free the instructor from the burden of remaining abreast of the latest developments in the various fields covered, but it means the book is likely to become rapidly dated; indeed, a few of the conclusions mentioned by the author have

already failed the test of time (for example, "applicability [of regional compensation models] to the continents appears dubious"). Not surprisingly, the quality of individual chapters seems to vary directly with the availability of good recent books on the subject. I was very impressed with the choice of topics and the discussion in the chapter on geomagnetism, much less so with the material on gravity, especially the section on the so-called gravity "inverse problem".

Worse, perhaps, topics from geodesy such as the hydrostatic figure of the Earth, Chandler wobble and precession of the equinoxes are missing, presumably due to the constraint of what can be covered in a one-semester course. This omission is understandable but unfortunate given the increasing influence of space-age geodesy on geodynamics, gravity and seismology. In addition, I found most of the homework problems predictable and uninspired compared to Stacey's inventive exercises.

Still, the book has its strong points. I suggest that instructors wishing to cover the global aspects of geophysics for students not yet comfortable with complex variables, spherical harmonics and Fourier analysis should give it careful consideration. □

Marcia McNutt is an Assistant Professor in the Department of Earth, Atmospheric and Planetary Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

Chemists at large

Peter Liss

Environmental Chemistry. By Peter O'Neill. Allen & Unwin: 1985. Pp.232. Hbk £20, \$30; pbk £8.95, \$14.95.

As Peter O'Neill remarks in the preface to this book, it is unfortunate that the terms "environmental chemistry" and "pollution" often go together. Environmental chemistry is much more than a study of the chemical effects of pollution. In order to assess how input of chemicals by anthropogenic activities can affect the environment, a proper understanding of how the system works in its unpolluted state is a vital requirement. The main aim of this book is thus to increase comprehension of how chemical processes operate in the environment. A few pollution topics are covered in their own right (for example, environmental problems involving lead, mercury, zinc and cadmium, as well as sewage-sludge disposal), and some others, such as acid rain, are used to illustrate the operation of fundamental chemical reactions in the environment.

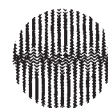
By and large the author is very successful in achieving his aim, and readers of the

book will certainly learn the elements of how the Earth operates as a natural chemical system. There are a few annoying lapses; for example, the discussion of sea-water buffer capacity is inadequate, and in dealing with the natural sulphur cycle too much emphasis is placed on hydrogen sulphide and too little on the role of dimethyl sulphide produced by algae near the ocean surface. On the other hand, the integration of concepts from pure chemistry into an environmental context is well done and organic aspects are particularly clearly presented.

Rather little knowledge of pure chemistry is assumed of the reader and this may limit the usefulness of the book for chemistry undergraduates. It should, however, have a ready audience amongst students studying biology, geology, environmental sciences, physical geography and ecology.

A glossary of terms used in the text is given at the end of the book. But it is a pity that no list of further reading is provided, because I am sure O'Neill's book will fire the imagination of at least some of its readers to study environmental chemistry in greater depth. □

Peter Liss is Professor in the School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK.



INRA
PUBLICATIONS

JUST PUBLISHED

**LES BESOINS
EN EAU
DES CULTURES
CROP WATER
REQUIREMENTS**

Conference internationale
Paris, 11-14 septembre 1984
L'Association Internationale
des Ingénieurs et de Chercheurs
C.I.T.

INRA

16 × 24, 927 p., 330 F

AVAILABLE IN APRIL 1986

INRA • OIE
PRECIS
DE
PATHOLOGIE
DES
POISSONS



21 × 28 358 p.

**MALADIES et
ACCIDENTS CULTURAUX
du TOURNESOL**

C. LAMARQUE



INRA

12 × 19 120 p.

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
SERVICE DES PUBLICATIONS
Rte de St-Cyr, F-78000 Versailles (France)
Télex : INRAPUB 699368F

Round and about the heterocycles

Kevin T. Potts

Heterocyclic Chemistry. By T.L. Gilchrist. Longman: 1985. Hbk £18.50, \$38.95; pbk £8.50.

OVER THE past decade increasingly sophisticated techniques have been brought to bear upon the synthesis of heterocyclic compounds, and upon the study of their properties and reactivity. Heterocyclic chemistry is no longer set apart from the main stream of organic chemistry, and organic chemists are finding many heterocyclic substrates to be important synthons. Dr Gilchrist, an active research worker in the field, has produced an excellent book. He builds on classical organic theory and provides a logical basis for understanding the varied reactions which heterocyclic systems undergo and the numerous strategies adopted for their synthesis.

Gilchrist has chosen his material well. He deals initially with the modern concepts of aromaticity and the criteria used to determine it, as well as with non-aromatic heterocycles where the consequences of bond angle strain in small- and large-ring heterocycles are clearly emphasized and torsional energy barriers are discussed for a variety of bonding situations.

In Chapter 4 the author develops a framework on which to build a rational approach to the vast array of synthetic procedures leading to heterocycles. The reactions involved in the ring-forming process are classified according to the synthetic method involved. Two broad groups of ring-forming reactions are considered: cyclizations, in which one bond is formed in the ring-closure process; and cycloadditions, in which two ring bonds are formed and in which no small molecules are eliminated. Although this scheme is an over-simplification well recognized by the author, it serves his purpose well in providing a systematic format to assist undergraduates and graduate students in assimilating a large amount of material.

Cyclization reactions are considered to involve intramolecular versions of the commonly encountered σ -bond-forming processes, especially the reaction of an electrophilic atom with a nucleophilic one. This group includes bielectrophilic reagents in which the electrophilic centres are separated by one or more atoms, as well as binucleophiles such as hydroxylamines (a 1,2-binucleophile), and others with the nucleophilic centres separated by one or more atoms. Reagents with electrophilic and nucleophilic centres also fit into this category, as do the syntheses of benzo-fused heterocycles employing ana-

logously substituted benzene derivatives. Used in illustration are examples of reactions involving displacements at saturated carbon, intramolecular nucleophilic addition to carbonyl groups and to other double bonds, cyclization onto triple bonds, radical cyclizations, carbene and nitrene cyclizations, and electrocyclic reactions. Numerous tables showing the reagents involved, possible intermediates and the products obtained complement and enlarge upon the discussion.

In keeping with the emphasis on cycloaddition reactions in the recent literature, Gilchrist focuses on their usefulness in leading to heterocycles with well-defined substitution patterns and, in many instances, with stereochemical control at more than one centre. Discussions of 1,3-dipolar cycloadditions, hetero-Diels-Alder reactions, [2+2] cycloadditions, chelotropic reactions and the Ene reaction are well illustrated by numerous examples and by tables which provide a clear summary of the salient features.

The remaining chapters (5–10) are a concise account of the synthesis and reactivity of all the commonly encountered three-, four-, five-, six- and seven-membered heterocycles containing one or more heteroatoms, as well as their benzo-

fused derivatives; the summaries of the modern chemistry of this family of heterocycles will whet the reader's appetite to learn more. Finally, there is a brief but helpful introduction to heterocyclic nomenclature.

The book appears at a particularly opportune time in that it provides an introduction to the more detailed reference collection, *Comprehensive Heterocyclic Chemistry*, published last year*. Problems are provided at the end of each chapter, the index is quite comprehensive, and altogether this will prove an ideal text for undergraduates and beginning graduate students. The more seasoned research chemist will also find much of value here, particularly as there are numerous references to both the original literature and review articles. □

Kevin T. Potts is a Professor in the Department of Chemistry, Rensselaer Polytechnic Institute, Troy, New York 12181, USA.

* For review see *Nature* **315**, 354 (1985). Pergamon have published a companion volume to this work, *Handbook of Heterocyclic Chemistry* by Alan R. Katritzky, a 542-page paperback based on the general chapters in *Comprehensive Heterocyclic Chemistry* and designed to give an overall view of the subject for students and researchers. Price is £22.95, \$29.95.

Behind mechanism

Neil S. Isaacs

The Physical Basis of Organic Chemistry. By Howard Maskill. Oxford University Press: 1985. Pp.490. Hbk £30, \$45; pbk £14.95, \$24.95.

Pictorial Orbital Theory. By John M. Tedder and Antony Nechvatal. Longman: 1985. Pp.120. Hbk £14.95, \$17.95; pbk £4.95.

PHYSICAL-organic chemistry is too large a subject to be encompassed in a single volume. Maskill's book aims to provide an account of the principles underlying techniques used routinely by physical-organic chemists in pursuit of reaction mechanisms. In this it fills a need, because many texts which discuss the evidence for accepted mechanisms cannot devote sufficient space to the theory on which the experimental evidence rests. The style and layout is pleasing, SI units are used throughout, and there are many worked examples (rather a lot in the gas-phase area) and tables of data. The author evidently expects a working knowledge of organic chemistry of at least second-year level in the reader.

The main strength of the book is to be found in the discussion of thermodynamic aspects of organic reactions and, although traditional in manner, the exposition is excellent. There is strong emphasis on

partition functions, with which organic chemists would be well-advised to become familiar. Chemical potential, for example, emerges nicely and activity coefficients are clearly defined. Acid-base equilibria receive a fairly standard but clear treatment, though the hard-soft principle suffers from the lack of perturbation MO theory, the basis of any explanation of this phenomenon; it is inadequate to discuss "hardness" in terms of polarizability (itself not clearly defined). The theory of rates is adequately expounded, as is the basic thermodynamics of catalysis, and the Eigen and Marcus theories and the Brønsted Law are well presented. Isotope effects are derived from partition functions, but more could be made of their mechanistic use, in particular the secondary β -effect. LFER are discussed, though not multi-correlation analysis or the resonance parameter, σ_R .

Regrettably, there are omissions. It has been the author's decision to exclude any discussion of molecular orbitals, to my mind a serious drawback, and there is little on the nature of intermolecular forces including solvation or the thermodynamic advantages of intramolecular processes. Conformational properties of ethanes are discussed; those of cyclohexanes but inadequately.

This is not, then, a new text on reaction mechanisms, nor does it encompass the entire "physical basis" of organic chemistry. Rather it will find a place as a helpful source of information on the basic physico-

al chemistry — kinetics and thermodynamics in particular — which is so extensively used and so often misunderstood by the organic chemist.

It is now customary to teach a considerable part of organic chemistry from the molecular orbital point of view. Simple Hückel theory is limited to the properties of π -systems and the inclusion of the MO properties of σ -bonded compounds has awaited the availability of reliable high-level approximations, such as the various *ab initio* programs now widely available. The MO parameters used in *Pictorial Orbital Theory* are taken from the collection published by Jorgensen and Salem in 1973.

Using qualitative arguments, Tedder and Nechvatal describe the building of MO's from AO's and the application to simple molecules such as HF, H₂O and methane. It is made clear that the two unshared pairs in water are of very different energies and symmetries, a fact often obscured by "rabbit ear" pictures. Frontier orbital pictures of larger molecules are then developed. Subsequent chapters deal with the use of frontier orbitals in the interpretation of some familiar reactions (the terms HOMO-gen and LUMO-gen, coined to replace "nucleophile" and "electrophile" will not meet with universal approval).

Radical abstractions are discussed in terms of HOMO-SOMO and LUMO-SOMO interactions, though mention of "polar factors" whose nature and origin are not explicit has meant the passages concerned are not as clear as they might be; perhaps the inclusion of atomic charges or of electrical potential contours would have been advantageous in this context. S_N2 reactions are shown to lead to inversion, and for E2-eliminations the necessity of LUMO lobes of the same sign spanning C_α and C_β is apparent. Additions of electrophiles to alkenes are obvious HOMO-LUMO interactions, though I harbour doubts as to the actual existence of Br⁺ in the reactions with bromine. Ambident reagents are discussed in the context of the hard-soft principle using Hückel MO's. The remainder of the book covers reactive intermediates, pericyclic reactions, benzenoid reactivity and aromaticity in a short and conventional treatment.

This will undoubtedly be a valuable book for all students of organic chemistry. Its particular strength lies in the inclusion of σ -bond reactions within the scope of MO theory, while the problems to be found in each chapter are very useful. On the negative side it suffers from being excessively brief, especially the later chapters which may be inadequate for an undergraduate course. □

Neil S. Isaacs is a Senior Lecturer in the Department of Chemistry, University of Reading, Whiteknights, Reading RG6 2AH, UK.

Compound relations from the metal-carbon bond

R.J.K. Taylor

Basic Organometallic Chemistry. By Ionel Haiduc and J.J. Zuckerman. *Walter de Gruyter*:1985. Pp.488. Hbk DM 169, \$64.90; pbk DM 79, \$29.90.

Metallo-organic Chemistry. By Anthony J. Pearson. *Wiley*:1985. Pp.398. Hbk \$45, £36; pbk \$15.90, £9.95.

THE publication of *Comprehensive Organometallic Chemistry* (Pergamon, 1982) marked a milestone in organometallic research. What is now required is an equally inspiring textbook on the subject. The task of writing such a book is daunting, however, given the wealth and diversity of material available for inclusion — organometallic chemistry encompasses the structure, bonding and physical properties of all compounds containing metal-carbon bonds, as well as their reactions and applications. In both of these texts, the authors have limited their horizons but, despite the similarity of the titles, they have concentrated on very different aspects of the field.

Basic Organometallic Chemistry was originally published in Romania in 1974 as *Chimia Compulsor Metalorganica*, with Haiduc alone as the author. This English-language version has been updated by the inclusion of results published up to 1984. It is divided into three parts, the first of which includes a discussion of the classification and bonding of organometallic complexes, a list of important literature sources and a brief section on laboratory techniques (although, surprisingly, no mention is made of syringe/septum methods). In Part II the compounds of the main group elements are dealt with group by group; those based on B, Si and As are included, whereas those based on P, S, Se and Te are not. Part III, which is subdivided according to ligand type, describes organometallic compounds derived from transition metals.

In general, the coverage of structural types in Parts II and III is up to date and comprehensive. However, the authors have concentrated on the synthesis and structures of organometallic complexes and have included little on their spectroscopic properties, reactivity and applications. In-text references are not provided, and although the hardback edition contains an extensive (111-page) bibliography this has been omitted from the paperback. The illustrations are generally clear, although opportunities are often missed to present valuable stereochemical information. More serious failings are the lack of

consistency in depicting bonding arrangements and the relatively large number of erroneous structures (including the formula accompanying the cover illustration).

Overall, then, Haiduc and Zuckerman have compiled a useful, up-to-date survey of organometallic compounds, but one which lacks sufficient information on reactivity and applications to stand alone as the text to support an undergraduate course.

In *Metallo-organic Chemistry*, Pearson limits himself to the chemistry of d-block organo-transition-metal complexes with particular reference to their applications in organic synthesis. The book is aimed at advanced undergraduates and organic graduates, but will also be of great benefit to practising organic chemists trying to keep abreast of this rapidly developing area of research. The first chapter deals with bonding in transition-metal-organometallic complexes, assuming a basic knowledge of frontier molecular orbital theory, while the second reviews typical properties and reactions of organometallic complexes, and includes sections on fluxionality, oxidative addition, reductive elimination, olefin metathesis, insertion reactions and catalytic processes. The remaining chapters then deal sequentially with the various ligand types (σ -alkyl, carbene and carbyne, η^2 , η^3 and so on).

No attempt has been made to provide a comprehensive catalogue of structural types; rather, attention has been concentrated on those complexes which are of most potential use in organic synthesis, and within this context the coverage is excellent and the examples well chosen. Observed reactivity patterns are clearly rationalized, the importance of spectroscopy (particularly NMR) is emphasized and synthetic applications are highlighted. Key references are given for each chapter, covering the literature up to 1982. The book is well illustrated, contains very few errors and a comprehensive index is included.

Metallo-organic Chemistry provides an excellent introduction to the subject from the organic chemist's standpoint. It should prove popular both for general reading and as a recommended text to accompany a lecture course. □

R.J.K. Taylor is a Lecturer in the School of Chemical Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Spring books supplement

The next review supplement to appear in *Nature* will be Spring Books, on 17 April.

Among the books to be reviewed are *Memoir of a Thinking Radish* (by P.B. Medawar), *Bird of Passage* (by Rudolf Peierls), *QED* (by Richard Feynman), *Unfinished Synthesis* (by Niles Eldredge), *The Case for Animal Experimentation* (by Michael Allen Fox) and *Problems of Biology* (by John Maynard Smith).

Third generation of organic chemistry

John Mann

Advanced Organic Chemistry: Reactions, Mechanisms and Structure, 3rd Edn. By J. March. Wiley:1985. Pp.1,362. \$39.95, £41.40.

Introduction to Organic Chemistry, 3rd Edn. By Andrew Streitwieser Jr and Clayton H. Heathcock. Macmillan/Collier Macmillan:1985. Pp.1,197. Hbk \$34, £40; pbk £17.95.

MARCH's *Advanced Organic Chemistry*, now in its third edition, has always been much more than a textbook. A fair indication of its scope can be gained from the facts stated on the dustjacket: 600 reactions considered in detail, and 10,000 references provided including 5,000 new to this edition. The book is designed for the advanced undergraduate or the post-graduate, though it commences with a good "revision course" in structure and bonding, reactive intermediates, acids and bases, and photochemistry. No mention is made of specialized topics such as primary and secondary metabolites, polymerization and so on, because the author believes these are best dealt with in monographs or review articles.

The vast bulk of the book, 870 pages in all, is thus devoted to the 600 reactions and their mechanisms. One novel feature in this edition is the use of IUPAC reaction names, and these are given in bold type; so, for example, hydrolysis of anhydrides becomes "hydroxy-de-acyloxy-substitution" and the Wittig reaction is an example of "alkylidene-de-oxo-bisubstitution". Mercifully, the old reaction names are provided in ordinary type.

The main strength of the book is the wealth of references. The Wittig reaction, for example, merits no fewer than 131 of them taken from the research literature or from books and reviews, together with 10 others to *Organic Syntheses*. It is the latter that are often the most useful, since they allow immediate access to tried and tested recipes for a particular reaction. At the end of the book there are two useful appendices, one giving information about use of the research literature, and the other providing a classification of reactions by type of compound synthesized,

thus allowing the reader to identify the relevant sections in the text. Comprehensive author and subject indices (170 pages in all) round things off.

The price of the hardback will deter many potential purchasers, and the publishers would do well to issue an international student edition. The book would then be widely used, and deservedly so.

For all its merits, *Advanced Organic Chemistry* lacks colour. That, according to Streitwieser and Heathcock in the preface to *Introduction to Organic Chemistry*, is what students need in order to "spotlight" key pieces of information. Again this is a third edition, and two-tone illustrations are a new feature of the book, bringing it into line with most of the standard textbooks on this subject.

There are now a large number of excellent organic texts for undergraduates, and

any choice between them must be largely subjective. This book has a particularly good chapter on NMR spectroscopy; a much expanded chapter on sulphur, phosphorus, and silicon chemistry; and molecular orbital theory now has its own chapter. One other useful feature is the inclusion of a number of "typical experiments" with experimental details, which provides a useful reminder to the student reader that chemistry is still an experimental science. Finally, a Student Study Guide (not seen by me), prepared by Paul A. Bartlett, is available, and if the preface is to be believed this should be almost as valuable as the main text, with its answers to problems, appendix of reactions and so on. □

John Mann is a Lecturer in the Department of Chemistry, University of Reading, Whiteknights, Reading RG6 2AH, UK.

Ways of industry

R.B. Wilcockson

The Chemical Industry. Edited by C.A. Heaton. Blackie: 1986. Pp.359. Pbk £15.95.

THE aim of this book, which complements a previous volume edited by Heaton (*An Introduction to Industrial Chemistry*, published by Blackie in 1984), is to provide a readable account of the principal sectors of the chemical industry. Taken together, the two volumes generally succeed well in covering this wide field concisely yet thoroughly. However their complementary nature is not readily apparent until one reads the introduction to the present volume, and it might have been better to use a common title and designate the books Volumes 1 and 2. The material in this book is set much better into context after reading the chapters "Sources of Chemicals", "The World's Major Chemical Industries" and "Petrochemicals" in its companion text (the rest of which deals with the more managerial and technology orientated aspects of the subject).

In the new book, six sectors of the chemical industry are covered, each by a different author, while a seventh chapter looks into the future more generally. Each account follows a nicely structured common format of brief history, present position, major products and the future. As in any such compilation, the chapters vary somewhat according to the background and approach of each author, as well as to the nature of the subject matter itself.

Chapter 1 ("Polymers") thoroughly covers processes and the importance of product properties, and helpfully explains the myriad abbreviations used as indus-

trial shorthand for the range of polymer products now available. Next comes an account of dyestuffs, with particular emphasis on chemical structures and synthesis routes, and then a concise but coherent chapter on a collection of inorganic chemical topics (the chlor-alkali, sulphur, nitrogen and phosphorus industries). Chapter 4 on pharmaceuticals concentrates largely on the extensive range of products and the influence of regulatory requirements on new product development; despite the diversity of pharmaceutical chemistry, this account would have benefited by inclusion of some synthesis routes with reference to starting materials and processing methods. Chapter 5 ("Agrochemicals") covers product types, environmental constraints and selected processes, and Chapter 6 the topical field of biological catalysis and biotechnology. The latter is an admirable review of a complex area in which the author wisely views the available and potential technologies as valuable tools rather than a panacea for all future processing needs.

Finally, the editor briefly discusses the current situation and picks out some trends for the future. Here, in considering computer applications, reference should perhaps have been made to the growing importance of automated batch-processing for speciality chemicals.

The books naturally lean towards British practices, but are set within a worldwide context. While a different arrangement of chapters would have helped each of the volumes stand better in its own right, together they should indeed prove useful both for students and for general reference, particularly in view of the educational trend towards a case-study approach to product technology. □

R.B. Wilcockson is a Lecturer in the Department of Chemical Engineering, Loughborough University of Technology, Loughborough LE11 3TU, UK.

Textbooks supplement — prices

Where possible both dollar prices in the United States and sterling prices in Britain are given in the bibliographical details for each book. Export prices will generally be higher. If a dollar or sterling price is not cited, the book is not available in one of the two countries. However readers should check both price and availability of books before ordering.

Choices for courses

Jon McCleverty

Inorganic Chemistry, 2nd Edn. By A.G. Sharpe. Longman:1986. Pp.696. Hbk £25; pbk £12.95.

ONE can regard the arrival of a new or revised textbook very much in the way one anticipates a full-course dinner. One knows it is likely to be filling, maybe nourishing, but will it be interesting, attractively presented, even memorable? Sharpe's second edition provides a repast very similar to his first, but with some refinement of selected courses.

The text is aimed at the early years of an undergraduate course in inorganic chemistry, and nearly everything of importance is included in some measure. The preliminaries deal with nuclear, electronic structure and theory, and physical properties of molecules. Solid state chemistry and the behaviour of inorganic compounds in aqueous and non-aqueous media are dealt with in three chapters. Altogether, one can regard the first eight chapters as a condensed diet of basic physical/"general" chemistry with heavy inorganic seasoning.

The main meal is structured around nine chapters dealing with the s- and p-block elements, presented in their vertical groups, and ten chapters on the d- and f-block elements. This second section is based mainly on a topic approach, for example, electronic structure and the consequences for bonding, and thermo-

dynamics, kinetics, metal carbonyls and organometallics. The d- and f-block elements are discussed in a limited way systematically, according to "tradition" (1st, 2nd/3rd row, lanthanides and actinides). In relation to the previous edition, there are additions to the discussions of donor-acceptor interactions, of the transition between non-metals and metals, and of aspects of polymer and bio-inorganic chemistry. Each chapter is referenced to general, more advanced, texts or specialist articles, and there are questions provided with numerical answers (where appropriate).

So what sort of rating does one give to this meal? In itself, it is satisfying dietarily and contains little extraneous or indigestible matter. When compared with other offerings, however, the rating is equivocal. Huheey, Mellor and Porterfield all use the same ingredients, but present them in different ways. There is a difference in philosophy in Sharpe's treatment — he deliberately accentuates thermodynamics rather than the symmetry, structure and bonding approach more common in the United States. Neither is preferable over the other, but in books of this size and price it is not possible to do justice to both. As far as presentation is concerned, the American texts look tastier. Sharpe's book is, however, interesting and memorable — anyone using this book in depth will find it clear and easy to read, if a trifle dull visually. □

Jon McCleverty is a Professor in the Department of Chemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK.

Chemistry in theory

Anthony Stone

The Chemical Bond, 2nd Edn. By J.N. Murrell, S.F.A. Kettle and J.M. Tedder. Wiley:1985. Pp.333. Hbk £19.95, \$31.85; pbk £8.95, \$14.30.

Structure and Spectra of Molecules. By W.G. Richards and P.R. Scott. Wiley:1985. Pp.172. Hbk £14.75, \$23.55; pbk £6.95, \$11.10.

Potential Energy Surfaces: Molecular Structure and Reaction Dynamics. By David M. Hirst. Taylor & Francis:1985. Pp.234. £19, \$38.

"MURRELL, Kettle and Tedder" has been a popular textbook for 20 years, in its first incarnation as *Valence Theory* and latterly as *The Chemical Bond*. The first edition of *Valence Theory* appeared in 1965, at a time when it was still a common belief that most chemists had no need of quantum mechanics, but being intended for theoretical specialists at the graduate level the book did not need to avoid mathematics. By the time the first edition of *The Chem-*

ical Bond appeared in 1978, it was usual for undergraduate chemistry courses to contain material on quantum mechanics and its application to chemical bonding. Regrettably, however, chemistry was then still widely thought to be a suitable subject for the semi-numerate, and many of the mathematical details had to be omitted.

For some years this deficiency could be made good by turning to *Valence Theory*. That is now out of print and is not even mentioned in the new edition of *The Chemical Bond*, which nevertheless remains very much the plain chemist's guide to the theory of chemical bonding; much of the mathematics is left out, although there is in the new edition a chapter on methods of *ab initio* calculation, introduced in the knowledge that programs to perform such calculations are now widely available.

It is a pity that in protecting the reader from the complications the authors have allowed some misleading material to remain, as for instance in the account of group theory. The treatment assumes that everything is real (without, as far as I can see, actually saying so). Unfortunately several of the statements of group-

theoretical results are incorrect in the complex case. It is possible, too, to disagree with the presentation in places. For example the discussion of Walsh diagrams is essentially an *explanation* of the behaviour of the calculated orbital energies of H₂O as the bond angle is varied, and the reader is not shown how to use the insights in Walsh's famous series of papers to understand and predict bond-angle relationships *without* performing elaborate calculations, or how to use descent of symmetry to identify the significant effects.

However these are minor criticisms of a wide-ranging, thorough and useful text. Richards and Scott's book *Structure and Spectra of Molecules* suffers from much greater deficiencies. It aims to emphasize the relationship between molecular structure and spectra, but cannot be said to succeed. The trouble is in part that a full appreciation of this relationship requires first a good understanding of the principles of molecular structure on the one hand and of spectroscopy on the other. However the account given here is extremely superficial. Formulae are plucked out of the air, without derivation; explanations are sketchy or entirely absent; and a sprinkling of typographical errors adds to the confusion.

It is notable that many important and valuable concepts have no objective reality. For example, we are accustomed to thinking of electronic structure in terms of molecular orbitals, but they are no more than theoretical constructs arising from a particular approximation. The same is true of the "potential energy surface" governing the motion of nuclei in molecules. Lying behind this idea, however, is the Born-Oppenheimer approximation, arguably the most successful approximation in molecular quantum mechanics. In much of chemical physics the potential energy surface provides the common ground between theory and experiment; it is a readily understandable concept which summarizes the results of *ab initio* calculations and provides the basis for the interpretation of experiments.

In his book, Hirst has sought to provide an introduction to the use of this fundamental idea and has succeeded rather well. He deals with the calculation of potential energy surfaces by both accurate and approximate methods, and their relationship to vibrational spectroscopy and reactive molecular beam scattering. As always in an introductory text one could wish for more detail in places; for instance, there is no quantitative discussion of the accuracy of the Born-Oppenheimer approximation away from intersections. On the whole, however, this is a careful account of an important topic. □

Anthony Stone is a Lecturer in the Department of Organic and Inorganic Chemistry, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, UK.

Physical refreshment

Maurice Rigby

Physical Chemistry, 3rd Edn. By P.W. Atkins. Oxford University Press/W.H. Freeman:1986. Pp.857. Hbk £35, \$37.95; pbk £14.95.

THE FIRST impression of the new edition of this best-selling textbook is that it is shorter than its immediate predecessor; indeed, it has almost 240 pages fewer. Closer inspection reveals that this has been achieved by the use of a changed page format, and comparison of equivalent passages suggests that any real difference in length is slight. The altered layout, with figures mostly in the wide margins, gives a fresh and attractive look to the material, which has been extensively revised.

Several new topics have been introduced in this edition, reflecting the continuing development of the subject. Applications of lasers are seen in discussions of laser isotope separation and of the Lamb dip technique (the spectroscopist's, not the sheep farmer's), and the phenomenon of oscillating reactions is now also described. The level both of the quantum mechanics and of the reaction kinetics has been increased, and several changes have been made to the sequence of presentation, most notably in the chapters on thermodynamics.

The latter section also sees the introduction of the recently recommended revision of thermodynamic standard states to a pressure of 1 bar, rather than 1 atmosphere. Though this has only a small or negligible effect on the values of tabulated thermodynamic functions and standard electrode potentials, it will require a change in teaching practice which will distress those who have yet to accept fully the use of SI units. The newly recommended symbol $\Delta_f H^\circ$ for 298.15K will take some time to be assimilated, as will the adoption of the IUPAC recommendation that we write $\Delta_f H^\circ$ rather than ΔH_f° . The reappearance of the concentration unit M in place of mol dm^{-3} , though not officially recommended, leads to a simplification of composite units, and will be welcomed by many.

Another new feature is the provision of sets of relatively straightforward applications of the principles and equations introduced in each chapter, on top of the wide range of problems found in earlier editions; a solutions manual, covering both of these collections of problems, is also available, and is a valuable complement to the main text. In addition, most of the longer tables of data formerly distributed throughout the book have been collected together in an appendix to give a useful data section of more than 20 pages.

No book describing such a range of material can hope to escape criticism completely, and some students may find the notational complexity of parts of the text to be rather intimidating. This, however, is a problem common to almost all degree-level books on physical chemistry, and forces one to wonder whether the use of a single comprehensive text is ideal, especially in the early stages of a course. This question apart, Atkins has again compiled an impressive survey of modern physical chemistry, one which seems certain to enjoy the popularity of the earlier editions. □

Maurice Rigby is a Lecturer in the Department of Chemistry, King's College London (KQC), Campden Hill Road, London W8 7AH, UK.

Science by numbers

Colin Upstill

Solving Equations with Physical Understanding. By John R. Acton and Patrick T. Squire. Adam Hilger: 1985. Pp.219. Hbk £24, \$36; pbk £12.95, \$19.

Introduction to Physical Mathematics. By P.G. Harper and D.L. Weaire. Cambridge University Press:1985. Pp.260. Hbk £20, \$42.50; pbk £6.95, \$14.95.

Partial Differential Equations for Scientists and Engineers, 3rd Edn. By G. Stephenson. Longman: 1985. Pp.161. Pbk £6.95, \$9.95.

Worked Examples in Mathematics for Scientists and Engineers. By G. Stephenson. Longman: 1985. Pp.239. Pbk £4.95, \$10.95.

MATHEMATICAL textbooks, even those aimed at scientists and engineers, used almost invariably to be impenetrable but impeccably rigorous and complete treatises, devoid of illustrations and written in a very formal style. Many still are. In some instances and for some students this is appropriate. More often than not, however, what is needed is less rigour and more explanation, less formality and more insight, and a lot of sketches. To the great benefit of today's students, more and more authors are aware of this, and are writing about mathematics for scientists and engineers without burying physical insight beneath a mountain of proofs of existence and the like.

Acton and Squire's *Solving Equations with Physical Understanding* embodies the notion that formal mathematics is an obstacle rather than an aid to understanding, and introduces and makes much of the "qualitative solution — trial function" method (QSTF). This feedback approach is fine in principle, but is dangerously over-reliant on solving by physical intuition, which works well only if your physical intuition is already keenly developed. Moreover the technique is dependent

upon everything being smooth and well-behaved, with a serious risk of missing interesting and important singular behaviour. As the authors admit, QSTF augments standard methods, it does not replace them. The risk is that the approach is only safe in the hands of those who know well enough what they are about not to need it.

Acton and Squire's wholly laudable emphasis on qualitative understanding and practical methods of solution is shared by the splendid *Introduction to Physical Mathematics* by Harper and Weaire. Written in a pleasant, informal style, in 34 short chapters, each ending with an excellent summary, it covers at a basic level many essential applications of mathematics in the physical sciences; the only serious flaw is a totally inadequate index. Not least among the book's delights is the account taken of the impact of modern computers.

This impact has been considerable on the solution of differential equations, and it is good to find a new edition of Stephenson's *Introduction to Partial Differential Equations for Science Students* repackaged as *Partial Differential Equations for Scientists and Engineers*, with an entirely new chapter on finite difference and finite element methods. A proper understanding of the underlying mathematics is essential if the common folly of applying numerical methods inappropriately is to be avoided. Such an understanding can be gained from this book, whose emphasis is unashamedly on methods of solution rather than the solutions themselves: it is a fairly formal text on linear partial differential equations, and the inclusion of numerical methods has given it a new lease of life.

The most effective and efficient way of learning scientific mathematics is not to struggle endlessly with a problem, but to consult a similar worked example. To this end, many textbooks include worked examples, and there are books which give a brief revision of theory followed by a wide selection of solved problems. Stephenson's *Worked Examples in Mathematics for Scientists and Engineers* is at the further end of this spectrum. With not even the briefest revision of theory, and with a complete lack of sub-headings, it is an indifferent example of the genre. Whilst it might be suitable for a student revising an entire course, it is much less useful to the reader who wishes to find an example of a particular sort; and, contrary to the assertion in the preface and on the cover, the book is almost useless for the postgraduate scientist wishing to recall his knowledge of a particular topic: he needs both methods and meaning. □

Colin Upstill is in the Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

Dynamical routes to quantum theory

William Rhodes

Quantum Mechanics: A Modern Introduction. By Ashok Das and Adrian C. Melissinos. *Gordon & Breach*:1985. Pp.640. Hbk \$65, £46; pbk \$29.50, £21.

The Picture Book of Quantum Mechanics. By Siegmund Brandt and Hans Dieter Dahmen. *Wiley*:1985. Pp.306. \$29.95, £26.50.

In *Quantum Mechanics*, Das and Melissinos have produced a novel and refreshing introduction to the fundamentals of quantum theory. Instead of the traditional development of the formal mathematical structure of the theory, which uses wave mechanics as a basis, the authors have adopted a more physical and operational approach. From the outset they emphasize the essential role of measurement in quantum mechanics, in contrast to classical mechanics, while such concepts as probability amplitudes, states and operators are formulated in terms of simple thought-experiments involving slit diffraction, particle spin and polarized light. The quantum nature of small systems is taken for granted from the beginning, rather than treating quantum mechanics as an (historical) outgrowth of classical mechanics.

No background in quantum theory is assumed. The discussion of each topic begins simply and progresses to an advanced, and often quite sophisticated, level. Dirac bra and ket notation is used throughout; accordingly, the coordinate and momentum representations, as well as discrete representations, are treated on equal footing. Many of the conventional methods, such as perturbation theory and the variational method, are introduced in an *ad hoc*, albeit complete and sound, manner. The structure of the theory is developed in a way which emphasizes the dynamical nature of physical systems and processes; that is, there is stress on the time-dependence of probability amplitudes for states and processes.

The book should be equally valuable to students of physics and chemistry. Simple physical systems of interest in both fields are mentioned freely, and often repeatedly, and the ammonia molecule is used as a model for several discussions including the one on symmetry. Figures are plentiful and adequate, and the thoughtful collection of problems will challenge the best of students. Chemists may find the system of units awkward, since it is more commonly used in high-energy physics, but this is no great handicap.

In places the writing is somewhat bland, but this is more than compensated for by

the simplicity and clarity of the authors' approach. My overall impression of the book is of its elegance and sophistication, and I anticipate it will find wide use in quantum theory courses.

In *The Picture Book of Quantum Mechanics*, Brandt and Dahmen present some of the elementary principles of quantum theory by means of computer graphics. As the title implies, pictures dominate the book, with the text providing description of the concepts and formulae for graphic representation. Both the elementary and the subtle features are depicted for such topics as single-particle wave packets moving in simple potentials, two-particle wave packets for distinguishable and indistinguishable particles, and scattering in three dimensions. The processes of tunnelling and the excitation and decay of metastable states are particularly well illustrated. There is a natural emphasis on the dynamical aspects of quantum mechanics.

This is a unique book. It does not provide a complete course in quantum theory, but as a companion work of reference it should be quite useful to students in providing insights into the dynamical structure of the theory. □

William Rhodes is Professor of Chemistry at Florida State University, Tallahassee, Florida 32306, USA.

Key appreciation

A.S. Dickinson

Computational Physics. By Steven E. Koonin. *Benjamin/Cummings*: 1986. Pp.409. \$32.50, £38.95.

THE advent of cheap microcomputers has given a fresh impetus to university-level physicists keen to include some practical experience of computing in their courses. Their objectives include enriching the undergraduate curriculum, introducing students to computational physics as a subject of growing importance and generally preparing them better for their future careers. Many lecturers can readily see

ways of including topics appropriate to their own specialism, but are rather daunted by the task of covering a large part of the whole course. Furthermore, preparing suitable material is very time-consuming.

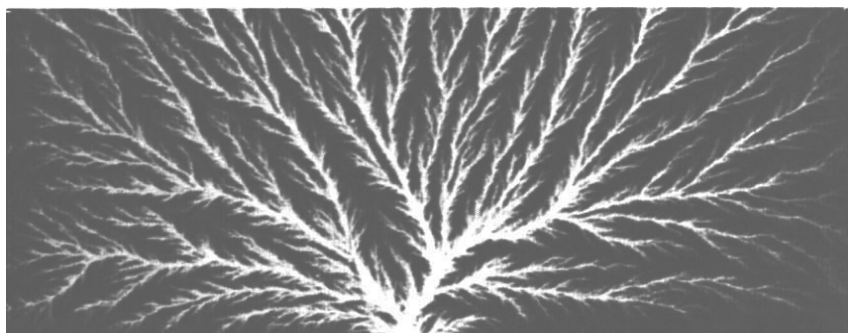
Such people will find this text of considerable value. A core set of numerical techniques is introduced and then illustrated by examples chosen to cover a wide range of topics, from classical potential scattering through self-consistent fields to two-dimensional fluid flow. The material is stated to be pitched at advanced undergraduate or beginning graduate student level. While some of the topics will be accessible to all students, several would probably be encountered only in optional courses (for example the eikonal approximation in quantum scattering, the schematic shell model in nuclear physics). For undergraduate teaching, a book with the physics at a rather lower level would be preferable, and would allow its introduction earlier in the course.

Complete BASIC listings of eight major examples and eight further projects are given for IBM and IBM-compatible machines, along with a five-and-a-quarter inch floppy disc of the source codes. (While this disc ran successfully on an IBM-PC, it failed on the only available IBM-compatible machine.)

One possible reservation concerns the introduction of numerical methods in the text rather than simply using off-the-shelf routines, but this is probably unavoidable when BASIC is used on a micro. Some readers may be annoyed by the use of non-SI units. And, as the author admits, with interpreted BASIC on the PC some codes run irritatingly slowly, so the compiled version is desirable.

Although the ultimate test of this text will be with students, I enjoyed sampling the exercises and projects; particularly effective use is made of graphical output. Overall, I welcome the book enthusiastically. As Koonin himself says, "it is perhaps best thought of as establishing an environment for further exploration". □

A.S. Dickinson is a Reader in the School of Physics, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK.



Pattern created by the electrical breakdown of a Plexiglas block subjected to a strong electric field. The picture is taken from Hans C. Ohanian's *Physics*, published by W.W. Norton. Price is \$36.95, £16.95.

Thermodynamics without tears

Tony Guénault

Thermodynamics and an Introduction to Thermostatistics, 2nd Edn. By Herbert B. Callen. Wiley: 1985. Pp. 493. \$45.15, £34.75.

The Theory of Thermodynamics. By J.R. Waldram. Cambridge University Press: 1985. Pp. 336. Hbk £30, \$59.50; pbk £10.95, \$24.95.

BY AND large, bed-time reading and thermodynamics do not go together. But here are two remarkably unrepresentative examples among the ensemble of thermodynamic textbooks.

Callen's classic *Thermodynamics* was first published in 1960, and well-thumbed copies are still to be found in every library. The fact that no revision has been necessary in the intervening 25 years speaks for itself, and the main change in this second edition is the inclusion of a compact section on thermostatistics (for want of a better word!). Callen's logic is that the macro- and micro-approaches to thermodynamics are not so distinct that a modern text should cover only one of them, but that the structure of either discipline should not be lost in the vague blanket coverage of some so-called thermal physics. The statistical method is introduced with admirable clarity through the treatment of an Einstein model of a solid. There are welcome additions on critical phenomena and quantum fluids, the discussions of irreversible thermodynamics and fluctuations have been reorganized, and many new and useful problems are included.

At first glance, the title of Waldram's book does not seem too appealing. However, "theory" here in no way means an obsession with computational detail, but rather a desire to expound clearly the underlying logical structure of the subject. Students are encouraged to investigate the necessary "applications" in problems at the end of each chapter. This is a most impressive piece of work — a thorough, scholarly and original contribution to the teaching of physics.

The unusual starting-point chosen by Waldram is a somewhat intuitive treatment of kinetics and the master equation. This then persuasively leads to expositions of statistical and classical thermodynamics, which are developed distinctly but from the same underlay. The whole text is beautifully graded, starting at a reasonably elementary (first-year university) level, but progressing in the later chapters to topics and ideas which, although immediately accessible to a physics graduate, are beyond the practical scope of

most third-year physics courses. An indication of the breadth is that among these topics are transport properties in metals, the law of mass action, imperfect gases and liquids, and graining and density matrices.

The clear structure of the book will enable it to be used for a variety of short courses, and at several levels. The reader who actually reads the preface and then the index will unearth three possible subsets of the coverage as a whole — one on classical thermodynamics, one comprising temperature, Boltzmann factors and kinetic theory, and one on chemical thermodynamics; incidentally, one wonders why a more accessible (diagrammatic?) method was not found for displaying this excellent feature. Otherwise the presenta-

tion is admirable in every respect; the pages are large and uncluttered, with useful marginal notes, and a complete index is provided.

These are two particularly well-structured and clearly written works. The stress of each of them is on the basic logic of the subject, each is satisfyingly successful and yet the solutions reached by the two authors are quite different. Any teacher or researcher will therefore have good reason for buying both books. But I suspect that many a student (with a pocket for only the one) will appreciate having Waldram's book to hand (by the bed, perhaps?) for years to come. □

Tony Guénault is a Reader in the Department of Physics, University of Lancaster, Lancaster LA1 4YB, UK.

Relative approaches

D.J. Raine

Basic Concepts in Relativity and Early Quantum Theory, 2nd Edn. By Robert Resnick and David Halliday. Wiley: 1985. Pp. 341. \$34.95, £36.95.

Elements of Relativity Theory. By D. F. Lawden. Wiley: 1985. Pp. 108. Pbk £4.95, \$7.95.

RELATIVITY and quantum theory provide interesting contrasts. Together they form the basis of modern physics, but the approach to the teaching of them has traditionally been very different.

With relativity one can go almost straight to the basic postulates, and the essential equations can be mastered with elementary algebra. But, despite references to "high gamma" engineering, there is not a lot to which one can practically apply the resulting knowledge of special relativity, on its own, at least. On the other hand, quantum theory is traditionally approached through several levels, some of which, like the Bohr theory and the more problematical aspects of wave-particle duality, have to be discarded before a complete understanding can be reached — if, indeed, the theory is completely comprehensible at all. Elementary introductions here tend to fall rather short in covering the abundance of realistic practical application. Furthermore, in teaching relativity the arguments historically used in the development of the theory can be employed. In approaching quantum theory the appropriate background in classical Hamiltonian mechanics cannot be assumed, and only the old quantum theory is available to bridge the gap.

The book by Resnick and Halliday divides along these traditional lines. What unifies it is the authors' clear vision of their audience, and an attention to detail

facilitated by a second edition. We are told that the authors have also written the most widely used physics textbook in history. I can quite believe it. Certainly, for its wide range of interesting problems and flexibility of organization one can forgive the occasional false *bonhomie* in references to Sally and Steve and their relatives; this book is an excellent text for a first-year introductory course.

Professor Lawden also knows his intended audience; his *Elements of Relativity Theory* grew out of lectures for science and engineering students not specializing in theoretical physics, and is now offered for more general consumption. The book gives a traditional, workmanlike account of the Lorentz transformations, using the k -calculus, and of relativistic mechanics, with occasional side-swipes at positivism (which he actually conflates with phenomenalism) and at popular mis-expositions. But it is indicative of the problems of this multi-level approach that an appendix deals with the derivative of x^2 (presumably in order to substantiate the claim that the only pre-requisite is elementary algebra), while a few pages earlier the geodesic equations have been written out.

Both books end with short introductions to general relativity which could scarcely be more different. In fact, it is now clear that the so-called geodesic principle, and the introduction of Riemannian geometry, are not features distinct from the equivalence principle; the gravitational redshift follows directly from the equivalence principle, not from the full theory, whereas the opposite is the case for the bending of light. And while the experimental tests of the theory may not be numerous, they are meagre only if here, too, one ignores recent developments. Needless to say, perhaps, Resnick and Halliday do not. □

D. J. Raine is a Lecturer in the Department of Astronomy, University of Leicester, Leicester LE1 7RH, UK.

Re-emergence in the solid state

Andrew K. Jonscher

Solid State Physics, 2nd Edn revised. By J.S. Blakemore. *Cambridge University Press*: 1985. Pp.506. Hbk £40, \$27.95; pbk £9.95.

SOLID state physics is a major branch of modern physics and is well served by textbooks at various levels of sophistication. Blakemore's text, which was first published in 1969 and again in 1974 by Saunders, may be regarded as one of the classics in this field, and I have long made use of it in undergraduate and introductory postgraduate teaching. It is good, then, to see the book back in print, especially because it is now available in Britain as a reasonably priced paperback and is thus within the financial reach of students.

This "revised" edition is only slightly

changed from the second edition of 1974. Essentially it is a corrected reprint, though the bibliography has been enlarged and a few other amendments have been made. However the book has stood the test of time reasonably well. It is eminently readable and, given the limitations of space and the breadth of possible subjects to be included, the choice of material remains well balanced.

The treatment does not include crystal-growing techniques and there is hardly a mention of p-n junctions, let alone of any other semiconductor devices. These are notable omissions. In the same vein there is nothing on quantum wells or two-dimensional electron gas effects, both of which are exciting and educationally valuable topics. One hopes that Blakemore will be persuaded to produce a fully up-to-date edition in the near future, but until that time the present book is well proven and altogether highly commendable. □

Andrew K. Jonscher is Professor of Solid State Electronics in the Department of Physics, King's College London (KQC), Pulton Place, London SW6 5PR, UK.

Thinking small

W.R. Gibson

Fundamental Particles: An Introduction to Quarks and Leptons. By Brian G. Duff. *Taylor & Francis*: 1985. Pp.122. Pbk £12, \$22.

Elementary Particles, 2nd Edn. By I.S. Hughes. *Cambridge University Press*: 1985. Pp.349. Hbk £27.50, \$49.50; pbk £9.95, \$19.95.

UNTIL a few years ago, particle physics was becoming more and more bogged down in detail. Now, both on the experimental and theoretical fronts, there have been enormous breakthroughs. From the experimentalists, using machines that caused particle-antiparticle annihilation to occur, came the discovery of the W^+ and Z^0 particles which confirmed theoretical ideas on the unification of two of the four fundamental forces in nature. The existence of gluons and quarks has been placed on a stronger footing and the quark-gluon concept has helped greatly in the interpretation of the available data. Theoretical work has led to a new, and most important, a simpler, understanding of the nature and the beginning of the Universe.

In the light of these successes, it is truly amazing that there are so few general books in this field, and that the implications for areas as various as religion, metaphysics, the existence of life and the Universe are discussed so rarely at a level appropriate for the layman and the student. Almost all the published material on

elementary particles is at a high academic level, describing complicated and *ad hoc* theories, and often with the main point lost in excessive detail. When the results could not be interpreted in a simple fashion it was hard to have it otherwise. But now the new experimental and theoretical results have blown away the cobwebs and opened up the prospect of the appearance of enlightening books on these topics.

Fundamental Particles by the late B.G. Duff falls into this category, and is a joy to read. It may not be accessible to all enquiring sixth-formers but at least it is a valiant attempt to reach them. While it is a short book, it lacks only more of the same material and fully captures all the current excitement. For that reason undergraduates and even seasoned researchers will find it well worth setting aside a couple of hours to read it.

Elementary Particles by I.S. Hughes is a second edition of a textbook written basically before the simplification became so apparent. An attempt has been made to update the contents with reasonable success, but the result is still a sober academic account of the field. This is a book for the serious student, similar to *Introduction to High Energy Physics* by D.H. Perkins, and is the ideal follow-up to Duff's *Fundamental Particles* which should have instilled in its readers the desire to learn in more detail about the fundamental advances that now characterize particle physics. □

W.R. Gibson is a Lecturer in the Department of Physics, Queen Mary College, University of London, Mile End Road, London E1 4NS, UK.

Waves of learning

W.G.V. Rosser

Electromagnetic Waves. By E.R. Dobbs. *Routledge & Kegan Paul*: 1985. Pp. 145. Pbk £4.50, \$9.95.

Maxwell's Equations and their Applications. By E.G. Thomas and A.J. Meadows. *Adam Hilger*: 1985. Pp.54. Pbk £2.95.

Problems and Solutions in Electromagnetic Theory. By C.M. Lerner. *Wiley*: 1985. Pp.614. Pbk \$34.95, £35.75.

Electromagnetic Wave Theory. By James R. Wait. *Harper & Row*: 1985. Pp.308. \$42.50, £49.

IN THEIR penultimate year as undergraduates, physics students generally take a long course on electromagnetism, using vector analysis and leading up to Maxwell's equations. These can then be taken as the starting point for work in the final year. There is much to be said for starting the final year with a series of lectures giving a broad overview of the main applications of Maxwell's equations, and Dobbs's *Electromagnetic Waves* provides an excellent choice of topics for a short course of this kind. The exposition is extremely clear and, even though many of the topics are fairly standard, they are generally discussed with fresh insights. Also included are problems for the student to solve. The book can be thoroughly recommended for the type of course for which it is written.

Maxwell's Equations and their Applications by Thomas and Meadows is restricted to 54 pages of text. It is a good idea to summarize the equations without long detours into topics such as AC theory. The stated aim of the book is "to explore the physical meaning of the equations and to bring out the relationship between this physical content and its mathematical representation".

The equation $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$ is developed from Coulomb's law. The magnetic pole concept is used to develop $\nabla \cdot \mathbf{B} = 0$ after the authors define "a magnetic intensity $\mathbf{B}$ corresponding to the electric intensity $\mathbf{E}$ ". This approach is not as popular as it once was. Commenting on the Lorentz force, Thomas and Meadows say that " $\mathbf{v} \times \mathbf{B}$ represents a kind of additional electric field which acts on a charged body in motion". Magnetostatics is developed from the Biot-Savart law in the differential form for a current element, which cannot be checked experimentally, while Ampère's circuital theorem is derived for the case of a long, straight wire only.

The equation $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$ is developed from experiments on electromagnetic induction, and, as an example of induction, on page 27 the authors consider "a cylindrical volume within which there is

a uniform $\mathbf{B}$ field", quoting the $\mathbf{B}$ field inside a long solenoid as an example. They assume that $\partial\mathbf{B}/\partial t$ is constant, claiming that, according to Maxwell's equations, the "time-varying $\mathbf{B}$ field gives rise to an $\mathbf{E}$ field", and adding "we have a static $\mathbf{E}$ field yet one that is not generated by charges". As an alternative "physical" interpretation, it is a straightforward boundary value problem to show that, inside the solenoid, in cylindrical coordinates the vector potential is $\mathbf{A}_\phi = \mu_0 n I r / 2$; then, using $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = (-\nabla\phi - \partial\mathbf{A}/\partial t)$ for quasi-stationary conditions we obtain $B_z = \mu_0 n I$ and $E_\phi = -(1/2)\mu_0 n r (dI/dt)$. It is straightforward to show that $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ for these fields. Hence both $\mathbf{B}$ and the induced electric field $\mathbf{E}$ can be determined *independently* from the vector potential and the equation $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ is just a relation between these fields. By integrating the relation $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$ at a fixed time we can equate $\oint \mathbf{E} \cdot d\mathbf{l}$ to the rate of change of magnetic flux, without relating $\mathbf{E}$ and $\mathbf{B}$ back to their common source, and without saying that the $\nabla \times \mathbf{E}$ produces the $-\partial\mathbf{B}/\partial t$ or vice versa.

When discussing the displacement current, Thomas and Meadows again adopt the old ether viewpoint, writing "a non-zero $\partial\mathbf{E}/\partial t$ gives rise to a spatially varying $\mathbf{B}$ "; yet, for example, the $\partial\mathbf{E}/\partial t$ term never appears as one of the sources of $\mathbf{A}$, and hence of $\mathbf{B} = \nabla \times \mathbf{A}$, in the expression for the retarded vector potential. Maxwell's equations are then used as the

basis of brief but lucid accounts of electromagnetic waves in empty space and conductors, and there are also good discussions of the Poynting vector and radiation pressure. Overall, this book only gives a brief and simplified development of the physical principles of Maxwell's equations and their applications, but students may find it a useful companion to the more detailed accounts of the subject.

Students should do a large number of problems to develop their expertise in electromagnetism. To this end, C.M. Lerner's *Problems and Solutions in Electromagnetic Theory* contains 425 solved problems covering all levels from first-year undergraduate to graduate; some problems are in cgs units, the others in SI units. The solutions are given in full, all the steps are explained clearly, and the book will be a very helpful complement to the standard textbooks on electromagnetism.

The central theme of *Electromagnetic Wave Theory* by J.R. Wait is the application of the impedance concept to wave phenomena. This approach has been particularly fruitful in ionospheric physics, antenna theory and geophysics, areas where Wait himself has made outstanding contributions. This is an advanced book which gives a comprehensive account of the methods concerned for the benefit of the more general reader at graduate level. □

W. G. V. Rosser is a Reader in the Department of Physics, University of Exeter, Exeter EX4 4QJ, UK.

Astrometric angles

Derek McNally

Spherical Astronomy. By Robin M. Green. Cambridge University Press: 1985. Pp.520. Hbk £40, \$69.50; pbk £15, \$27.95.

AT A time when astrometry is assuming an increasingly prominent role in astronomical research, it is essential to have available a good, up-to-date undergraduate textbook on the subject. R.M. Green was responsible for the last (1977) revision of Smart's classical *Spherical Astronomy*, but while there is a kinship between that book and this new one published under the same title, the content has been changed entirely.

A feature of textbooks such as Smart's and my own *Positional Astronomy* of 1974 was the exclusion of mathematical techniques which would not be familiar to beginning students. Both Smart and myself also concentrated on providing explanation of simplified practices which would allow rapid determination of parameters of limited accuracy. However, considerable advances in astrometric precision have

been made in recent years, and more are likely in the next decade with the advent of the astrometric satellite HIPPARCOS and the Space Telescope. While a rigorous relativistic treatment has been given by Murray in *Vectorial Astrometry* (1983), there has remained a need for a book intermediate in level between Smart/McNally and Murray. That need is well met by the present text.

The book has the advantage that it gives a simpler introduction than Murray to the relativistic basis for astrometric measurement. It also retains a derivation of the older, first-order formulae for day-to-day use. It is important, as it always was, that an astrometric textbook lays the foundation for a thorough development of the subject, while recognizing that only a few of its readers will ever want to use the full detail.

Dr Green develops his arguments vectorially, which is a concise and useful way of working and provides a direct link to Murray's monograph. He explains the recent changes in the form of dynamical time used in the *Astronomical Almanac* and also gives a short summary of the mechanics of precession as well as the more usual description of its effects. It is



Andromeda Galaxy, photographed through the 0.9m telescope of the Lick Observatory. The picture is taken from the paperback edition of *The New Astronomy*, by Nigel Henbest and Michael Marten, to be published next week by Cambridge University Press. Price is £9.95, \$14.95.

pleasing to note that precession is treated here in terms of three rotations, rather than by the algebraically simpler method used by Smart. A very welcome new addition to a book at undergraduate level is an account of the astrometry of radioastronomy; this is a field where astrometric results of the highest precision are being obtained and it is likely to play a leading role in future work on fundamental reference systems.

There are a few niggles. For example, more use might have been made of vector methods. On the other hand, while a text of this sort is under continual tension between the need for clarity and the need for conciseness, the lack of an extended account of the calendar, seasons, rising, setting and twilights is regrettable. It is surprising, too, to see a discussion of dependences under plate measurement — dependences are really only justified as a quick laboratory technique giving equatorial or rectangular equatorial co-ordinates directly.

Altogether, though, Dr Green has produced a good standard work, pitched at about second-year undergraduate level, which sets out the modern basis and terminology of astrometry. One hopes it will lead to a much better understanding, appreciation and use of astrometry by future generations of astronomers. Many of today's astronomers remain unaware of the contributions, actual and potential, of the subject — they now have little excuse for not ensuring that their successors are astrometrically aware. □

Derek McNally is a Senior Lecturer in the Department of Physics and Astronomy, University College London, and Assistant Director of the University of London Observatory, Mill Hill Park, London NW7 2QS, UK.

All at sea — and up in the air

Henry Charnock

Elements of Dynamic Oceanography. By David Tolmazin. *Allen & Unwin*: 1985. Pp.181. Hbk £25, \$39.95; pbk £11.95, \$19.95.

Introductory Physics of the Atmosphere and Ocean. By L. Hasse and F. Dobson. *Reidel*: 1986. Pp.126. Dfl. 57, \$19.95, £15.95.

THE complicated motion of our turbulent, stratified, rotating (and ill-observed) ocean presents problems among the most difficult in fluid mechanics. To deal with them in words and pictures, without mathematical formalism, is a formidable task. But Professor Tolmazin has to a large extent succeeded in communicating his enthusiasm about ocean circulation with only few equations. Perhaps his book should have been subtitled, like one of Auden's, *The Romantic Iconography of the Sea*: the first words of Tolmazin's preface are "The ocean evokes the most romantic images of nature".

So far as there is a unifying image in the book, it is the Stommel paradigm originally associated with explaining the westward intensification of ocean boundary currents: Tolmazin deals with the general meridional flows due to the curl of the windstress, with currents such as the Gulf Stream, with the analogous western boundary currents at great depth and with the meso-scale eddies that have been such a feature of recent studies. But there is much more in his book than a simplified account of the development of ocean circulation research over the past 50 years — there are sections on turbulence (including Kolmogoroff's theory of locally isotropic turbulence) and on recent developments in observing techniques such as satellite altimetry and acoustic tomography, as well as descriptions of more classical instruments, observations and interpretations.

In places the reader is expected to take too much on faith (in deriving the Richardson number, for example, is it really "natural to assume that the kinetic energy of a particle is proportional to its density and the vertical velocity gradient"?), and the explanatory comments sometimes seem unnecessarily picturesque (do trains in the Northern Hemisphere really wear the right-hand rail faster than the left?). Some of the background material is not quite up to date (the definition of salinity, for example), and there are some minor errors and a number of misprints, but the book generally succeeds in communicating the fascination of a rapidly developing subject. Moreover it

will be of interest to established Western oceanographers because it has many references to Soviet research; the author was until a few years ago an active researcher in the Soviet Union, as will be seen from some of his illustrations and from his occasional use of an unusual word or sentence.

Oddly enough, Tolmazin's basically descriptive book would be more appropriately entitled *Introductory Physics of the Ocean*, while those of Hasse and Dobson (there are two essentially separate sections bound together) would be more accurately described as *Elements of the Physics of the Atmosphere and Ocean*. Both Hasse and Dobson start from the basic building blocks of meteorology and physical oceanography, and go on to provide much-condensed accounts of atmospheric and oceanic processes, assuming that the reader is familiar with the

mathematics and physics involved. For example, the equations of motion, in tensor notation, are presented in Hasse's first few pages while the last 16 of Dobson's part of the book are devoted to an expert summary of the dynamics of sea surface waves. This is because both sections were originally presented as introductory material to a conference of physicists and chemists concerned with air/sea transfer processes.

It seems unlikely that these two accounts will be accessible to readers not already familiar with the subject. However they may well provide a useful reference summary for fluid dynamicists and physicists seeking an indication of how their specialities can be applied to the atmosphere and the ocean. □

Henry Charnock is a Professor in the Department of Oceanography, University of Southampton, Southampton SO9 5NH, UK.

Climatology snapped

T.M.L. Wigley

World Climatic Systems. By John G. Lockwood. *Edward Arnold*: 1985. Pp.292. Pbk £17.50, \$29.50.

CLIMATOLOGY has become like the Hydra, the creature itself being the climate system and each head, in Lockwood's terminology, being a subsystem. Lockwood has taken on the Heracleian labour of attempting to describe the beast by photographing all the heads at once. This is a difficult task, now more than ever because the subject is a target moving with increasing speed. Is the present book a success? Certainly most of the heads have been, if not completely captured, at least quite interestingly portrayed.

The book is directed towards second- and third-year university and polytechnic students in geography, environmental science and related subjects such as agriculture and hydrology (the latter field, surprisingly, is not mentioned in the preface, even though it is a major component of the text). It is divided into two parts, a large section on the climate systems and a smaller one on climatic impacts. The first part considers the atmospheric, oceanic, cryospheric and land-surface subsystems separately (with the land-surface further divided into arid, grassland and forest subsystems). The second unfortunately seems to have been rather an afterthought, and covers only climate and energy, and climate and food (ignoring the impact of climate on water resource (sub)systems, although this topic is covered to some degree in various places in the climatic system section).

The structure of the book is sensible and attractive. The execution is less good,

simply because of the breadth of material that must be covered. Consequently, the writing is very dense and at times superficial, especially in the impact section. Much of the content is too obviously derivative ("according to..." is one of the author's favourite phrases) and, as a result, the cited material tends to be reviewed uncritically. On the other hand, the coverage is impressive and the references are extensive and reasonably up to date (although published in 1985, most of the sources mentioned come from the years up to and including 1983).

There are a few errors of fact, most of them of no importance, but still reprehensible. More annoying to the climatological community are the omissions; very few of the priorities identified by the World Climate Programme are covered well, if at all, while too much space is devoted to Lockwood's own particular interest in hydrological runoff modelling. The balance of the book would therefore appeal less to those climatologists whose background is in meteorology. When compared, for example, with *The Global Climate* (edited by J.T. Houghton and published by Cambridge University Press in 1984 and again in paperback last year) one finds little overlap, although this is not necessarily a bad thing.

This is not an inspiring book like Schneider and Londer's *The Coevolution of Climate and Life* (see *Nature* 310, 517; 1984), but it is a valuable, workmanlike addition to the climatological literature. As a "snapshot", it provides a reasonable balance between scope and resolution and should give students a good feeling for both the breadth of disciplines that climatology now embraces, and the complexities within each of them. □

T.M.L. Wigley is Director of the Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.